Pramana – journal of physics

Editor

R Nityananda
Raman Research Institute, Bangalore

Associate Editor

H R Krishnamurthy
Indian Institute of Science, Bangalore

Editorial Board

G S Agarwal, *Physical Research Laboratory, Ahmedabad*
V Balakrishnan, *Indian Institute of Technology, Madras*
J K Bhattacharjee, *Indian Assoc. for the Cultivation of Science, Calcutta*
D N Bose, *Indian Institute of Technology, Kharagpur*
B M Deb, *Panjab University, Chandigarh*
Rohini M Godbole, *Indian Institute of Science, Bangalore*
S Kailas, *Bhabha Atomic Research Centre, Bombay*
R K Kaul, *The Institute of Mathematical Sciences, Madras*
A V Khare, *Institute of Physics, Bhubaneswar*
T Padmanabhan, *Inter-Univ. Centre for Astronomy and Astrophysics, Pune*
R Ramaswamy, *Jawaharlal Nehru University, New Delhi*
A K Raychaudhuri, *Indian Institute of Science, Bangalore*
K C Rustagi, *Centre for Advanced Technology, Indore*
E V Sampathkumaran, *Tata Institute of Fundamental Research, Bombay*
Abhijit Sen, *Institute for Plasma Research, Gandhinagar*
S K Sikka, *Bhabha Atomic Research Centre, Bombay*
Y Singh, *Banaras Hindu University, Varanasi*

Editor of Publications of the Academy

V K Gaur
C-MMACS, NAL, Bangalore

Annual Subscription Rates

All countries except India (Price includes AIR MAIL charges)	US$ 200
India	Rs. 200

Annual subscriptions for **individuals** for India and abroad are Rs. 100/- and $50 respectively.

All correspondence regarding subscription should be addressed to the **Circulation Department** of the Academy

---

Editorial Office

Indian Academy of Sciences, C V Raman Avenue,
P.B. No. 8005, Bangalore 560 080, India

Telephone: 334 2546
Telefax: 91-80-334 6094

Email: pramana@ias.ernet.in

© 1997 by the Indian Academy of Sciences. All rights reserved.

"Information for contributors" is printed in the last issue of every volume.

Pramana – journal of physics

Volume 49
1997

Published by
Indian Academy of Sciences
Bangalore 560 080

Pramana – journal of physics

Editor

R Nityananda
Raman Research Institute, Bangalore

Associate Editor

H R Krishnamurthy
Indian Institute of Science, Bangalore

Editorial Board

G S Agarwal, *Physical Research Laboratory, Ahmedabad*
V Balakrishnan, *Indian Institute of Technology, Madras*
J K Bhattacharjee, *Indian Assoc. for the Cultivation of Science, Calcutta*
D N Bose, *Indian Institute of Technology, Kharagpur*
B M Deb, *Panjab University, Chandigarh*
Rohini M Godbole, *Indian Institute of Science, Bangalore*
S Kailas, *Bhabha Atomic Research Centre, Bombay*
R K Kaul, *The Institute of Mathematical Sciences, Madras*
A V Khare, *Institute of Physics, Bhubaneswar*
T Padmanabhan, *Inter-Univ. Centre for Astronomy and Astrophysics, Pune*
R Ramaswamy, *Jawaharlal Nehru University, New Delhi*
A K Raychaudhuri, *Indian Institute of Science, Bangalore*
K C Rustagi, *Centre for Advanced Technology, Indore*
E V Sampathkumaran, *Tata Institute of Fundamental Research, Bombay*
Abhijit Sen, *Institute for Plasma Research, Gandhinagar*
S K Sikka, *Bhabha Atomic Research Centre, Bombay*
Y Singh, *Banaras Hindu University, Varanasi*

Editor of Publications of the Academy

V K Gaur
C-MMACS, NAL, Bangalore

Annual Subscription Rates

All countries except India (Price includes AIR MAIL charges)	US$ 200
India	Rs. 200

Annual subscriptions for **individuals** for India and abroad are Rs. 100/- and $50 respectively.

All correspondence regarding subscription should be addressed to the **Circulation Department** of the Academy

---

Editorial Office

Indian Academy of Sciences, C V Raman Avenue,
P.B. No. 8005, Bangalore 560 080, India

Telephone: 334 2546
Telefax: 91-80-334 6094

Email: pramana@ias.ernet.in

© 1997 by the Indian Academy of Sciences. All rights reserved.

"Information for contributors" is printed in the last issue of every volume.

Pramana – journal of physics

CONTENTS – VOLUME 49

(July–December 1997)

Number 1

Special Issue: Conference on Fundamentals of Physics & Astrophysics

Editorial Note 1

Preface 3

Quantum Mechanics

Causal quantum mechanics . . . . . *Virendra Singh* 5

A generalized Schrödinger formalism as a Hilbert space representation of a generalized Liouville equation . . . . . *R K Varma* 17

The quantum geometric phase as a transformation invariant . . . . . *N Mukunda* 33

Supersymmetry in quantum mechanics . . . . . *Avinash Khare* 41

Violation of signal locality and unitarity in a merger of quantum mechanics and general relativity . . . . . *Partha Ghose* 65

Particle Physics

Quark confinement in 2 + 1 dimensional pure Yang–Mills theory . . . . . *Spenta R Wadia* 71

CP symmetry and its violation in fundamental interactions . . . . . *S D Rindani* 81

Plasma Physics

Chaos and its implications . . . . . *B Buti* 93

Nonlinear wave modulations in plasma . . . . . *N N Rao* 109

Thermodynamics of nonlinear plasma systems . . . . . *P K Kaw* 129

Astrophysics

The large scale structure of the universe: Dynamical and statistical aspects . . . . . *Varun Sahni* 131

Rotation curves of galaxies: Missing mass or missing physics . . . . . *V Krishan* 147

Statistical Physics

Deposition and evaporation of k-mers: Dynamics of a many-state system . . . . . *Mustansir Barma* 155

Number 2

Review

Cosmological N-body simulations. . . . . *J S Bagla and T Padmanabhan* 161

Research Articles

Differentiability of invariants and Liapunov equations in Hamiltonian systems . . . . . *B R Sitaram* 193

Truncated harmonic oscillator and parasupersymmetric quantum mechanics . . . . . *B Bagchi, S N Biswas, Avinash Khare and P K Roy* 199

A new class of cosmological models in Lyra geometry . . . . . *G P Singh and Kalyani Desikan* 205

Inhomogeneous cosmological models with heat flux . . . . . *L K Patel, Ramesh Tikekar and Naresh Dadhich* 213

Photon production from baryon rich quark-gluon plasma . . . . . *Pragya, Ashok Goyal and J D Anand* 225

Some two body non-leptonic B decays and estimation of f_D in the heavy quark effective theory . . . . . *A K Giri and R Mohanta* 231

Modified WKB transmission for fusion . . . . . *S V S Sastry and S K Kataria* 239

Excitation functions of alpha particle induced reactions in aluminium and copper . . . . . *S Mukherjee, B Bindu Kumar and N L Singh* 253

Number 3

Projection method for 4d non-orthogonal hyperlattices . . . . . *S Jayanthi, R K Mandal and S Lele* 263

Four constant potential function of amorphous selenium and evaluation of its thermodynamic and acoustic properties . . . . . *R Venkatesh, R V Gopala Rao and Sumita Bandyopadhyay* 269

High pressure study of phase transitions in α-$FePO_4$ . . . . . *Chitra Murli, Surinder M Sharma, S K Kulshreshtha and S K Sikka* 285

Thermoelectric power of high-T_c superconductors using extended Hubbard model . . . . . *S Keshri and P K Barhai* 293

Radially dependent antishielding factor $\gamma(r)$ of the Sc^{3+} ion in the solid state . . . . . *B B Panigrahi, N C Mohapatra and S Hafizuddin* 301

Time dependent magnetization of a $Zn_{0.1}Fe_{0.9}Fe_2O_4$ fine magnetic particle system . . . . . *R V Upadhyay* 309

Fine structure in the time of flight distribution of C_2 in laser produced plasma from graphite . . . . . *S S Harilal, Riju C Issac, C V Bindhu, V P N Nampoori and C P G Vallabhan* 317

Number 4

Review

Some aspects of Kaluza–Klein cosmology . . . . . *S K Srivastava* 323

Research Articles

The algebra and geometry of $SU(3)$ matrices . . . . . *K S Mallesh and N Mukunda* 371

Accurate energy-levels of the sextic-double-well oscillator $V(X) = X^6 - 3X^4$ using modified Hill determinant approach . . . . . *Biswanath Rath* 385

Hopping of chaotic dynamics in a doubly resonant parametric oscillator . . . . . *Jolly Jose, Bimalendu Deb and S Dutta Gupta* 389

The effect of environment induced pure decoherence on the generalized Jaynes–Cummings model . . . . . *Gautam Gangopadhyay and S H Lin* 399

Isothermal fluid sphere: Uniqueness and conformal mapping . . . . . *Naresh Dadhich* 417

Investigation of the boundary value problem corresponding to a generalized Skyrme lagrangian . . . . . *Rajesh Kumar and P K Chattopadhyay* 421

Electromagnetic properties of a chiral-plasma medium . . . . . *H Torres-Silva, C Villarroel, P H Sakanaka and N Reggiani* 431

Bulk viscosity of neutron stars . . . . . *V K Gupta, A Wadhwa, S Singh and J D Anand* 443

Number 5

Application of density-functional perturbation theory to calculate nonlinear polarizabilities of helium-like systems . . . . . *Arup Banerjee and Manoj K Harbola* 455

Polarization and dichroism in atomic Auger spectroscopy following absorption of circularly polarized light . . . . . *S Sen* 461

Two Coulomb waves theory for direct excitation . . . . . *Sima Chakrabarti and Durga P Dewangan* 479

On the water-induced critical double point in a polymer solution: The role of isotope substitution . . . . . *B M Jaffar Ali, S Venkatachalam and A Kumar* 493

Self-consistent numerical study of pure and impurity doped three-band Hubbard model . . . . . *Prasenjit Sen and Avinash Singh* 505

Measurement of fast neutron induced fission cross-section of thorium using Lexan plastic track detector . . . . . *R K Jain, H S Virk, J Rama Rao and S K Bose* 515

SCORPION: A system for coincidences between recoil and projectile ions at NSC, New Delhi. . . . . *M J Singh, S K Goel, R Shanker, D O Kataria, N Madhavan, P Sugathan, J J Das, D K Avasthi and A K Sinha* 521

TDPAC measurements in textured hcp samples . . . . . *R Govindaraj and K P Gopinathan* 535

Brief Reports

Phonon dispersion in aluminium arsenide and antimonide . . . . . *Prafulla K Jha, Sujata Rath and Sankar P Sanyal* 547

Slowness surfaces, ray velocity surfaces and phonon focussing in high T_c superconductor crystals. . . . . *J Philip and M S Kala* 555

Number 6

Darboux transformation and elementary exact solutions of the Schrödinger equation. . . . . *Vladislav G Bagrov and Boris F Samsonov* 563

Detection of determinism and randomness in time series: A method based on phase space overlap of attractors . . . . . *M S Gopinathan* 581

Schrödinger picture formalism of Φ^6 theory. . . . . *K P Satheesh and K Babu Joseph* 591

Motion of test particles around domain walls. . . . . *Subenoy Chakraborty and Lalit Biswas* 603

Einstein pseudotensor and total energy of the universe . . . . . *N Banerjee and S Sen* 609

Semi-empirical formula for Λ-binding energies in ground states of light hypernuclei . . . . . *M Z Rahman Khan, Nasra Neelofer and M A Suhail* 617

Measurement of excitation functions and mean projected recoil ranges of nuclei in ^{12}C-induced reactions on vanadium. . . . . *M Ismail, R P Sharma and M H Rashid* 623

Brief Reports

Spherical aberration and its correction using Lie algebraic methods . . . . . *Govindan Rangarajan and Minita Sachidanand* 635

Two body non-leptonic Λ_b decays in the large N_c limit. . . . . *A K Giri, L Maharana and R Mohanta* 645

Editorial Note

Pramana, in the recent past, has published the proceedings of selected specialized meetings in the area of particle physics (December 1993 and October 1995) and also a substantial collection of invited articles in the area of nonlinear science (January–February 1997). It has sometimes been suggested that we could also carry the proceedings of meetings which have a different, broader character. One such was held in 1995 at the Physical Research Laboratory in Ahmedabad. Since many of the participants felt that the meeting successfully brought together stimulating ideas from a wide range of topics in physics, we are attempting, as an experiment, to convey the original flavour, even when it involves elements of speculation and controversy. This is part of an ongoing process of finding new ways in which *Pramana* can be useful to its readers, and ultimately their reactions will guide our future efforts.

Rajaram Nityananda
Editor

Preface

A small meeting of distinguished physicists working in different fields was organized at PRL from March 14 to March 16, 1995. The broad objective was to bring together experts who would raise some fundamental questions in their field and bring their own perspective to the discussion. The choice of topics reflects to a large extent the interests of PRL scientists. Consequently, the meeting was called "Fundamentals of Physics and Astrophysics". Before the meeting, we had no intention of bringing out the proceedings. At the conclusion, however, it was felt by all the participants that it would be worthwhile to have a published record. We are therefore very thankful to all the speakers who agreed to provide a manuscript of their talk. This took some time but finally we received all of them.

The subject matter of the talks is divided into five categories: i) Quantum Mechanics, ii) Particle Physics, iii) Plasma Physics, iv) Astrophysics, v) Statistical Physics.

In the first category V Singh discusses the measurement problem in Quantum Mechanics in the context of separation of system from the measuring apparatus. This led to the development of Causal Quantum Mechanics. R K Varma presented his work on formulating a conceptually unified description of Classical and Quantum Mechanics. These two talks were in the realm of foundations of Quantum Mechanics. The remaining three contributions in this category relate to other developments in Quantum Mechanics. N Mukunda talked about a kinematic approach to describe the Berry phase. Avinash Khare discussed the development of Supersymmetry in Non-relativistic Quantum Mechanics. Some of the difficulties that arise in a merger of Quantum Mechanics and General Relativity were brought out by Partha Ghose in his talk.

In Particle Physics, the talks were addressed to two important questions. Spenta Wadia discussed the question of confinement in 2+1 dimensional Yang Mills theory. The CP symmetry and its violation were the subject of S D Rindani's talk.

There were three talks on Plasma Physics. B Buti gave a general account of chaos and its implication in Plasma Physics. N N Rao reviewed many aspects of nonlinear wave modulations in plasmas. P K Kaw talked about thermodynamics of nonlinear plasma system. (These proceedings carry a reference to the complete paper published elsewhere).

In Astrophysics, V Sahni described Large Scale Structure of the Universe. V Krishan talked about the Rotation Curves of Galaxies and the problem of missing mass.

In the Statistical Physics section, there is one contribution by M Barma. For a model complex system in which there are many steady states he showed how to find a labelling scheme for the large number of steady states. This problem is of crucial importance in glass, spin glass and rubber.

We believe the contributions contained in this volume would be of value to the physics community particularly because the current trend is towards highly specialized meetings.

We would also like to express our thanks to Mr V T Viswanathan for very extensive help first with typing several manuscripts and making all of them computer ready.

A C Das
S B Khadkikar
J C Parikh

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 5–16

Causal quantum mechanics

VIRENDRA SINGH
Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400 005, India

Abstract. We discuss the split between system and measuring apparatus, i.e. non-unitary nature, of the conventional quantum mechanics to motivate a causal unitary description of nature. We then describe causal quantum mechanics of de Broglie and Bohm. We conclude by presenting a version of recently proposed causal quantum mechanics which treats position and momentum variable symmetrically.

Keywords. Causal quantum mechanics; de Broglie–Bohm theory.

PACS No. 03.65

1. Is the world classical and/or quantum?

1.1 *Classical world*

The description of the physical universe given by classical physics was internally consistent and complete. It was a realistic conception of the world. The classical physics describes the world as it is. The physical world was made up of two kinds of entities viz. particles and waves. They obeyed well defined dynamical laws of change. It was thus a deterministic world. The only role of probability in classical physics, e.g. in statistical mechanics, was due to our lack either of knowledge or of interest in the initial conditions. It was not basic. Another important feature of classical physics was that it satisfied the principle of 'Einstein-locality' i.e. the factual situation of a system S_1 does not change due to whatever measurements may be made on a system S_2, which is spatially separated from S_1.

There was no split in classical physics between system which is being studied and the apparatus with which we study it. Both were described by the classical physics. The interaction which coupled the system and the apparatus was a normal physical interaction. By using the gentle, and if needed even gentler, probe it was possible to learn about the properties of the system under study to any desired degree of accuracy. What one learnt this way were the properties of the system as they were. There were thus no epistemological problems associated with a measurement.

The description of the world in classical physics is thus unitary without a split between System-Apparatus. Even though this is a very satisfying feature of classical physics description of nature, the classical physics failed in providing a useful description of atomic phenomenon and quantum mechanics had to be introduced by founding fathers viz. Planck, Einstein, Bohr, de-Broglie, Heisenberg and Schrödinger in the first quarter of this century.

Do we continue to have a unitary picture of the physical universe in quantum physics as was the case in classical physics? The answer is no, in the standard way quantum mechanics is formulated. By standard way we refer here to Copenhagen interpretation of quantum mechanics, and since this is not monolithic, we also include its variants. We now proceed to elaborate the negative answer using Bohr's and von-Newmann's views about measurement in quantum mechanics.

1.2 *Bohr*

According to Bohr [1] the system of interest, which we are studying and which is generally microscopic, must be described as a quantum object using laws of quantum physics. However the measuring apparatus, which is generally macroscopic, must be described as a classical object obeying laws of classical physics as the language of classical physics is the only language in which we can convey the results of our experiments. When we study a system S with experimental arrangement E what we observe is not just a property of S as it was but various outcomes depend on the total setup S and E. Quantum physics provides us with a calculus of probabilities for various outcomes of the joint setup S and E. Unlike classical physics there is no way in which E dependence can be wished away.

We must emphasis that in Bohr's view the classical physics has logical priority in as much as we cannot do physics without it. This logical priority is different from chronological priority, which of course is there, of classical physics over quantum physics. Thus physical universe is classical and quantum.

These views of Bohr raise a number of problems. How is one to understand the working of an apparatus? Should it follow from classical physics or should we use quantum physics when we want to do that. In that case apparatus is a quantum system when we want to understand it but is a classical system when we want to use it as a measuring device. What is the interaction between a quantum system and its classical measuring system? Well, there is no articulated theory of that. Then there is the question of how much of a given 'system + apparatus' is to be put on the quantum side and how much on the classical side. In practice this may not cause much problem as we believe that if we will put enough on the quantum side we will achieve the required degree of accuracy. This is supposed to be ensured by the folkloric 'correspondence principle' according to which quantum mechanics should, in appropriate limits, go into classical mechanics. But in principle the split remains.

1.3 *von Neumann*

In contrast to Bohr it was decided by von Neumann [2] to treat both the system S and measuring apparatus A quantum mechanically. Let the eigenfunctions of the system S corresponding to an observable Ω, be denoted by ψ_m, i.e. $\Omega\psi_m = \omega_m\psi_m$, with an eigenvalue ω_m. Further let the pointer readings a be the eigenvalues of the 'pointer position operator' M, for the system A with the eigenfunctions $f(a)$ i.e. $Mf(a) = af(a)$.

Consider first system S to be in an eigenstate ψ_n of Ω. According to von Neumann the 'measurement interaction' causes the system-apparatus state $\psi_n f(a)$ to become $\psi_n f(a_n)$.

From the pointer reading a_n we conclude thou the system was in the state ψ_n of the observable Ω with the eigenvalue ω_n. We now go over to case where the system S is not in a pure state of Ω but is in a linear superposition given by $\psi = \sum_n C_n \psi_n$. By the linearity of quantum evolution equation, the measurement interaction would cause the system-apparatus state $(\sum_n C_n \psi_n) f(a)$ to develop into $\sum_n C_n \psi_n f(a_n)$. From this we can, of course, conclude that the probability of measurement resulting in the pointer reading a_n, implying that the system was in the eigenstate ψ_m with eigenvalue ω_m is given by $|C_m|^2$. But note that this wave function still does not correspond to a definite pointer reading for the apparatus. The apparatus is still in a linear combination of states with different, even macroscopically different, pointer readings i.e. in a grotesque state.

von Neumann postulates at this stage the infamous 'collapse of the wave-function', according to which, when the measurement is completed and the pointer reading is a_n (say), the wave function $\sum C_n \psi_n f(a_n)$ collapses to $\psi_n f(a_n)$. He calls this a process of the first kind. These are acausal, discontinuous and takes place only when the measurement is completed and observation registered. In an elaboration by London and Bauer [3], and subscribed to by Wigner [4], the measurement is completed only at the level of human consciousness of the result of the measurement. In contrast to these processes of first kind there is the usual process of second kind describing the evolution of the system between measurements. These are causal, continuous and described by Schrödinger equation.

Thus irrespective of whether we follow Bohr, where quantum physics has to be supplemented by 'classical apparatus' or von Neumann, where we have to postulate a causal processes of the first kind (wave function collapse), we find that either way quantum physics is inherently incomplete. It is inexactly formulated. Quantum physics does not provide a unitary description of the nature.

1.4 *Quantum mechanics for a closed system*

This inexactness in the fundamentals of the quantum mechanics for a theory, as basic as this theory, is intolerable and there have been various attempts to formulate a quantum mechanics for a closed system. Another motivation has been the inability of any formulation of quantum mechanics, with an artificial split between system and apparatus, to deal with the quantum mechanics of the whole universe i.e. quantum cosmology. Wheeler's phrase "Include the observer in the wave function" expresses the same desire. Some of these attempts are

(i) Causal interpretation of de Broglie [5] and Bohm [6]. This was also favoured by Bell [7] and others [8].

(ii) Many world interpretation suggested by Everetts, initially championed by Wheeler, and followed later by de Witt and others [9].

(iii) 'Quantum mechanics of the history' approach and related concept of decoherence (Griffith, Gell-Mann and Hartle, Omnes, Zurck etc.) [10].

In the next § 2, we shall describe de Broglie–Bohm causal interpretation of the quantum mechanics restricting ourselves to the nonrelativistic quantum mechanics. We also point out some aspects of this theory which make us look for a version of it treating position and momentum symmetrically. This is described in § 3 mainly in the context of

nonrelativistic quantum mechanics of one nonrelativistic particle moving in one space dimension but with some remarks on situations without these restrictions.

2. de Broglie–Bohm causal interpretation of quantum mechanics

2.1 *A bit of history*

Louis de Broglie had proposed a 'pilot wave theory' interpretation of quantum mechanics in 1927 which provided a realistic interpretation of quantum mechanics [5] unlike the pragmatic one associated with Copenhagen School. Under the criticism of Pauli and others, and lack of endorsement even from those physicists, such as Einstein, who did not like Copenhagen interpretation, the interpretation did not gain currency. Even de Broglie gave it up. An important role was also played in this by so called 'von Neumann' theorem [2] which was taken to mean by physics community that a completion of quantum mechanics with extra (hidden?) variables was not possible if it had to reproduce the observable predictions of quantum mechanics.

Bohm, in 1952, independently came up with a similar proposal [6]. He was able to take care of various objections which had been earlier raised by Pauli, de Broglie and others. Besides he supplemented it with his ideas on theory of measurement.

It is a historical accident that de Broglie–Bohm theory was almost totally ignored until quite recently [11]. Bell's work showing irrelevance of 'von-Neumann theorem' to hidden variable program, and his advocacy of this theory, as an explicit counter example to Copenhagen orthodoxy's claims on nonexistence of a causal description for quantum physics, has had a lot to do with the revival of interest in it.

2.2 *Mathematical formulation of de Broglie–Bohm theory*

Consider first a single particle, with mass m, moving in the potential $V(\mathbf{r})$. The time evolution of this system is described by the Schrödinger equation for the wave function $\psi = \psi(\mathbf{r}, t)$,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi. \tag{2.2.1}$$

Let us use the polar decomposition

$$\psi = R\,\mathrm{e}^{iS/\hbar} \tag{2.2.2}$$

with R and S real. Substituting (2.2.2) in eq. (2.2.1) and separating real and imaginary parts of the equation we obtain [12]

$$\frac{\partial R}{\partial t} + \frac{1}{2m}[R\nabla^2 S + 2\nabla R\cdot\nabla S] = 0 \tag{2.2.3}$$

and

$$\frac{\partial S}{\partial t} + \left[\frac{(\nabla S)^2}{2m} + V(x) - \frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}\right] = 0. \tag{2.2.4}$$

Defining

$$P_{(\mathbf{r},t)} = |\psi|^2 = [R(\mathbf{r},t)]^2, \tag{2.2.5}$$

$$m\mathbf{v}(\mathbf{r},t) = \nabla S(\mathbf{r},t) \tag{2.2.6}$$

and

$$U(\mathbf{r},t) = -\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}, \tag{2.2.7}$$

we obtain from (2.2.3) and (2.2.4) the more suggestive equations (2.2.8) and (2.2.9) given below. We have

$$\frac{\partial P}{\partial t} + \nabla(P\mathbf{v}) = 0 \tag{2.2.8}$$

which is of the form of a continuity equation for position probability density P with an associated current density $P\mathbf{v}$. We also obtain

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + U = 0, \tag{2.2.9}$$

which is of the form of the classical Hamilton–Jacobi equation for a particle of mass m moving in a potential $V + U$ i.e. sum of the given external potential V and a quantum potential U.

We have so far only manipulated the Schrödinger equation mathematically into the suggestive forms (2.2.8) and (2.2.9). We now come to de Broglie–Bohm causal version of the quantum mechanics.

We associate with this system a wave function $\psi(\mathbf{r},t)$ obeying Schrödinger equation and a trajectory function $\mathbf{q}(t)$ of the particle. The trajectory function $\mathbf{q}(t)$ is the additional hidden variable introduced so as to be able to give a causal description. The particle trajectories are given by the dynamical equation

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \mathbf{v}(\mathbf{q},t) = \frac{1}{m}\nabla S(\mathbf{r},t)\,|_{\mathbf{r}=\mathbf{q}}\,. \tag{2.2.10}$$

The particle moves in a potential $V + U$. Note that this description is nonlocal since through U the particle experiences a potential which depends on the whole set up as the wave function $\psi(\mathbf{r},t)$ is determined by the conditions over the whole space through Schrödinger equation. Further even if wave function is small in some region, i.e. the probability $P = R^2$ of its being there is small, its influence in that region need not be small as the quantum potential

$$U = -\frac{\hbar^2}{2m}\cdot\frac{\nabla^2 R}{R}$$

may not be small there.

In this theory the motion of the quantum particle is analogous to the motion of boat which is set on an automatic pilot guided by radio wave signals. What matters for this purpose is the signal, i.e. the shape of the wave, and not its intensity. The boat moves under its own energy. The wave function ψ plays the same role for directing the motion of

for the de Broglie–Bohm theory.

The N-particle generalization is straightforward. One now supplements the N-particle wave function $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N st)$ with N trajectory functions $\mathbf{q}_1(t), \ldots, \mathbf{q}_N(t)$ obeying the dynamical equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t} q_i = \nabla_i S|_{\mathbf{r}_i = \mathbf{q}_i}, \tag{2.2.11}$$

where

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t) = \mathrm{Re}^{iS/\hbar} \tag{2.2.12}$$

and the quantum potential is given by

$$U = -\frac{\hbar^2}{2m} (\nabla_1^2 R + \nabla_2^2 R + \cdots + \nabla_N^2 R)/R|_{\mathbf{r}_i = \mathbf{q}_i(t)} . \tag{2.2.13}$$

One, of course, also has N-particle Schrödinger equation.

2.3 *Probability in de Broglie–Bohm theory*

It would be remarked that all the equations of de Broglie–Bohm theory, given in § 2.2, are deterministic. Given $q(t)$ and $\psi(\mathbf{r}, t)$ at $t = 0$ we can calculate them for later times. So how do quantum probabilities arise? They arise here in the same way as they do in classical statistical mechanics i.e. through our ignorance of initial conditions.

We are not able to control precise position of the particles. All we are able to prepare are statistical ensembles at a given time t, say $t = 0$, in which the position variable q is distributed with probability function $P(\mathbf{q}, t = 0)$ given by

$$P(\mathbf{q}, t = 0) = |\psi(\mathbf{q}, t = 0)|^2. \tag{2.3.1}$$

Once one has done that it can be shown that de Broglie–Bohm dynamics is such that we have

$$P(\mathbf{q}, t) = |\psi(\mathbf{q}, t)|^2. \tag{2.3.2}$$

We thus see that, in view of this result, the probability distribution (2.3.1) is somewhat special.

2.4 *Ontology*

In the de Broglie–Bohm interpretation a quantum object, say an electron, has both a particle aspect, it has a trajectory $q(t)$ associated with it, as well as a wave aspect, as the wave function ψ is involved in determining its velocity at any time. This is the resolution of the usual wave particle duality conundrums of the standard Copenhagen interpretation in this theory. It is not a wave or a particle depending on the measuring setup but it is always a wave and a particle. For example in the usual two slit experiment the wave always goes through both the slits but the particle goes through only one of them. The trajectories of the particles, however, are not straight lines, as they normally are in classical mechanics for free particles, due to the presence of quantum potential. The wave

function, and therefore, the quantum potential, does depend on the presence of both the slits even if the particle goes through only one of them.

Particle position is a hidden variable of the theory as we cannot control its value and there is inability, at least at present, to have any other distribution than $|\psi(\mathbf{q}, t)|^2$. Particle position, in this theory, however, has a special role. It 'is intrinsic and not inherently dependent.... on the overall context' [13]. Firstly, it is conceptually independent of the wave function with its own dynamical equation of motion. Secondly, it can be measured without being changed. In the terminology of Bell it is a 'beable' [7].

Particle momentum p is postulated to be given by an expression equal to mass times particle velocity i.e.

$$p_i = \nabla_i S \equiv (p_i)_{dBB} \tag{2.4.1}$$

and is also a hidden variable of the theory. This clearly depends on the phase of the wave function of the system as a whole. The momentum, unlike position, is not regarded as an intrinsic property of the particle. A measurement does not reveal a momentum value as given by (2.4.1).

The phase space density $\rho(p, q; t)$ is taken in this theory to be given by

$$\rho(p, q; t) = |\psi(q, t)|^2 \delta(p - (p)_{dBB}). \tag{2.4.2}$$

Note that, as expected

$$\int \rho(p, q; t) \mathrm{d}p = |\psi(q, t)|^2, \tag{2.4.3}$$

but

$$\int \rho(p, q, t) \mathrm{d}q = \delta(p - (p)_{dBB}) \tag{2.4.4}$$

and

$$\neq |\tilde{\psi}(p, t)|^2, \tag{2.4.5}$$

the momentum space density, where $\tilde{\psi}(p, t)$ is the momentum space wave function of the system.

3. Causal quantum mechanics with a symmetrical treatment of position and momentum

3.1 *Asymmetry in de Boglie–Bohm theory between position and momentum*

There is a clear asymmetry in the treatment of particle position and momentum variables in the de Broglie–Bohm theory as noted earlier. This aspect of the theory was considered as a defect of the theory [14].

Takabayashi pointed out while the phase space density $\rho(p, q; t)$ correctly reproduced the phase space average of a general function $f(q)$ of position q i.e.

$$\int \rho(p, q; t) \mathrm{d}p f(q) = \langle \psi \mid f(q) \mid \psi \rangle = \int \mathrm{d}q \mid \psi(q) \mid^2 f(q) \tag{3.1.1}$$

in view of (2.4.3), it does not do so for an arbitrary function $g(p)$ of momentum p i.e.

$$\int \rho(p,q;t)\mathrm{d}q\, g(p) \neq \langle\psi \mid g(p) \mid \psi\rangle = \int \mathrm{d}p\, |\tilde{\psi}(p)|^2\, g(p) \tag{3.1.2}$$

in view of (2.4.4) except when $g(p)$ is a linear function of momentum [15]. In view of this

$$(\delta\mathbf{q})^2 = \langle\psi|\, q^2|\psi\rangle - |\, \langle\psi|\mathbf{q}|\psi\rangle\,|^2$$

is reproduced while

$$(\delta p)^2 = \langle\psi|\mathbf{p}^2|\psi\rangle - |\, \langle\psi|\mathbf{p}|\psi\rangle\,|^2$$

is not and as such the product $\delta p\ \delta q$ involved in Heisenberg uncertainty principle is not recovered.

We may further note that, even for a 'free' particle, i.e. particle not acted on by external potentials, the momentum $(p_i)_{dBB}$ does not satisfy the 'expected' equation of motion

$$\frac{\mathrm{d}(\mathbf{p})_{dBB}}{\mathrm{d}t} = 0.$$

Why should position variable be special? One argument in favor of this position given is that ultimately all observations can be reduced to 'pointer readings'. At least it looks like that in our normal experience despite position variable having no preferred role in standard quantum mechanics as compared to, say, momentum [16]. This could, however, a quirk of nonrelativistic mechanics. In Galilean group it is possible to define a position variable which is conjugate to momentum. For Poincaré group such is not the case. In fact for massless particles such as photons there is no generally acceptable position variable. In contrast for light we have direct comprehension that the given light is say green i.e. of its wavelength which is related to its momentum. So in relativistic S-matrix theory and zero mass particle it would seem that it is momentum variable which is special and not the position variable.

3.2 *A symmetrical treatment for the case one particle moving on a line*

We now propose a reformulation of causal quantum mechanics with a symmetrical treatment of particle position and momentum [17]. For this purpose our point of departure is to replace (2.4.1) by

$$p = \hat{p}(q,t),$$

where $\hat{p}(q,t) \neq (p)_{dBB} = \nabla S$, and is to be determined by the requirement that the phase space density $\rho(p,q;t)$ reproduces correct marginal distributions. Let the phase space density ρ be given by the ansatz

$$\rho(p,q;t) = |\psi(q,t)|^2\, \delta(p - \hat{p}(q,t)). \tag{3.2.1}$$

The requirement about the marginal distributions is given by

(i) $$\int \rho(p,q;t)\mathrm{d}p = |\psi(q,t)|^2, \tag{3.2.2}$$

(ii) $$\int \rho(p,q;t)\mathrm{d}q = |\tilde{\psi}(p,t)|^2\,. \tag{3.2.3}$$

The ansatz (3.2.1) satisfies (3.2.2) automatically. The requirement (3.2.3) leads to the equation for the determination of $\hat{p}(q,t)$ given by

$$\int_{-\infty}^{\epsilon q} \mathrm{d}x' \,|\psi(\epsilon x';t)|^2 = \int_{-\infty}^{\hat{p}(q,t)} \mathrm{d}p' \,|\tilde{\psi}(p',t)|^2, \tag{3.2.4}$$

where $\varepsilon = +1$ or -1 represents a twofold ambiguity in the determination of $\hat{p}(q,t)$.

We also require that phase space density satisfies Liouville equatiuon i.e.

(iii) $$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho \dot{q})}{\partial q} + \frac{\partial(\rho \dot{p})}{\partial p} = 0. \tag{3.2.5}$$

This leads to the dynamical equation of motion

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \left.\frac{\mathrm{d}S(x,t)}{\mathrm{d}x}\right|_{x=q} \tag{3.2.6}$$

which is the same relation as in the de Broglie–Bohm causal quantum mechanics, by integrating (3.2.5) over the variable p and comparing with the continuity equation for $|\psi(q,t)|^2$. We further demand that the dynamics in the phase space is Hamiltonian i.e. there exists a Hamiltonian $H_c(p,q)$ such that

(iv) $$\dot{q} = \frac{\partial H_c}{\partial p}, \quad \dot{p} = -\frac{\partial H_c}{\partial q}. \tag{3.2.7}$$

The Hamiltonian H_c is not to be confused with the Hamiltonian H which is involved in the Schrödinger equation

$$i\frac{\partial \psi}{\partial t} = H\psi. \tag{3.2.8}$$

The requirement (3.2.5) and (3.2.7) are satisfied provided we choose

$$H_c = \frac{(p - A(x))^2}{2m} + V(x,t) \tag{3.2.9}$$

for the case when

$$H_{QM} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + U(x). \tag{3.2.10}$$

Here A is given by

$$\hat{p}(q,t) = (p)_{dBB} + A(q,t), \tag{3.2.11}$$

and there is an appropriate equation to determine $V(x,t)$.

Illustrative examples for a quantum free particle and oscillator can be found in ref. [17].

The potential felt by the particle in de Broglie–Bohm theory was external potential plus a quantum potential. For us, in view of (3.2.11), the momentum is not just a product of mass and particle velocity but acquires, like the potential, a quantum correction $A(q,t)$. An additional contribution like this to momentum occurs even classically in the presence of electromagnetic fields.

3.3 *Some features of the symmetrical treatment*

The expression (2.4.2), while actually symmetrical between position and momentum, is not manifestly so. It can however be put in such a form given by

$$\rho(q,p;t)=|\psi(q,t)|^2|\tilde{\psi}(p,t)|^2\delta\left(\int_{-\infty}^{\varepsilon q}\mathrm{d}q'\,|\psi(\varepsilon q',t)|^2-\int_{-\infty}^{\hat{p}(q,t)}\mathrm{d}p'\,|\tilde{\psi}(p',t)|^2\right). \tag{3.3.1}$$

It is obviously positive definite i.e.

$$\rho\geq 0 \tag{3.3.2}$$

unlike the Wigner phase space density. It reproduces marginal position and momentum distributions correctly.

The class of functions for which phase space density ρ reproduces the correct expectation value now is $f(q)+g(p)$, i.e.

$$\int \mathrm{d}p\,\mathrm{d}q\rho(p,q;t)[f(q)+g(p)]=\langle\psi\,|\,f(q)+g(p)|\psi\rangle \tag{3.3.3}$$

while this was restricted to only

$$f(q)+cp$$

in de Broglie–Bohm version.

We further note that the symmetrical version also leads to Newton's law

$$\frac{\mathrm{d}\hat{p}}{\mathrm{d}t}=0 \tag{3.3.4}$$

for a free particle (i.e. $U=0$ in (3.2.10)).

We may also note that despite this treatment lacking manifest Galilean invariance, it is in fact Galilean invariant [18].

3.4 *Generalizations*

If the space dimension is greater than one and/or if the number of particles is more than one, then the number of marginal distributions is no longer two as was the case for one particle in one dimension.

Consider for example two particles moving in one space dimension. Let the position and momenta of the two particles be denoted by (q_1,p_1) and (q_2,p_2). We may prefer to use centre of mass and relative positions and momenta given by

$$P=p_1+p_2,\quad p=p_1-p_2,\quad Q=q_1+q_2,\quad q=q_1-q_2.$$

We now have four marginal distributions for the joint probability densities for the pairs of variables given by (P,p), (P,q), (q,Q) and (p,Q). It turns out that it is not possible to find a positive definite $\rho(P,p,Q,q;t)$ such that all the four marginal distributions are reproduced. It is however possible to reproduce marginal joint distributions in the first

three pairs, such that

$$\dot{P} = 0 \text{ (Newton's law)},$$

$$\dot{q} = \frac{\partial S}{\partial q}$$

and another equation for $\dot{Q}$.

Some results are known now as to what marginal distribution cannot be reproduced [19]. Further work is however needed to elucidate the general structure underlying these considerations.

Acknowledgement

The author wishes to thank Prof. S M Roy for many discussions on the foundations of quantum mechanics.

References

[1] N Bohr, *Atomic theory and the description of nature* (Cambridge, 1934); *Atomic Physics and Human Knowledge* (Wiley, New York, 1958); *Essays 1958/1962 on Atomic Physics and Human Knowledge* (Wiley, New York, 1963)
L D Landau and E M Lifshitz, *Quantum mechanics* (Pergamon, 1958)

[2] J von Neumann, *Mathematische Grundlagen der Quanten Mechenik* (Springer-Verlag, Berlin, 1932) [*Mathematical Foundations of Quantum Mechanics*, translated by R T Beyer (Princeton, 1955)]

[3] F London and E Bauer, *La Theorie de l'observation en Mecanique Quantique* (Hernan and Cie, Paris, 1939). [An English translation is available in *Quantum Theory and Measurement* edited by J A Wheeler and W H Zurek (Princeton, 1983)]

[4] E Wigner, *Symmetries and reflections* (Indian Univ. Press, Bloomington, 1967)

[5] L de Broglie, *Compt. Rend.* **183**, 447 (1926); **184**, 273 (1927); **185**, 380 (1927)
L de Broglie, *Nonlinear wave mechanics, A causal interpretation* (Elsevier, 1960)

[6] D Bohm, *Phys. Rev.* **85**, 166 (1952); **85**, 180 (1952)
D Bohm and B J Hiley, *The undivided universe: An ontological interpretation of quantum theory* (Routledge, London, 1993)

[7] J S Bell, *Speakable and unspeakable in quantum mechanics* (Cambridge, 1987)

[8] P R Holland, *The quantum theory of motion* (Cambridge, 1993)

[9] H Everett III, *Rev. Mod. Phys.* **29**, 454 (1957)
B S Dewitt and N Graham (eds) *The many world interpretation of quantum mechanics* (Princeton, 1973)

[10] R B Griffiths, *J. Stat. Phys.* **36**, 219 (1984); *Am. J. Phys.* **55**, 11 (1987)
M Gell Mann and J B Hartle, in *Proc. of the Third International Conference on the Foundations of Quantum Mechanics in the light of New Technology* edited by S Kobayashi, H Ezawa, Y Murayama and S Nomura (Tokyo, 1990)
J B Hartle, *Vistas in Astronomy* **37**, 569 (1993) and references therein
R Omnes, *Rev. Mod. Phys.* **64**, 339 (1992)
W H Zurek, *Phys. Rev.* **D24**, 1516 (1981); **D26**, 1862 (1982)

[11] J T Cushing, *Quantum mechanics: Historical contingency and Copenhagen hegemony* (Chicago, 1994)

[12] E Madelung, *Z. F. Physik* **40**, 332 (1926)

[13] D Bohm and B J Hiley, *The undivided universe: An ontological interpretation of quantum theory* (Routledge, London, 1993) p. 109

[14] See e.g., W Pauli, in *Louis de Broglie, Physicien et Penseur* edited by A George (Paris, 195
[15] T Takabayashi, *Prog. Theor. Phys.* **8**, 143 (1952)
[16] E J Squires, *Found. Phys. Lett.* **3**, 87 (1990)
D Home, *Phys. Lett.* **A190**, 353 (1994)
[17] S M Roy and V Singh, *Mod. Phys. Lett.* **A10**, 709 (1995); *Proc. of the Conf. on Bose a Twentieth Century Physics* (Calcutta, 1994)
[18] A Kumar, *Int. J. Mod. Phys.* (to appear)
[19] A Martin and S M Roy, *Phys. Lett.* **B350**, 66 (1995)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 17–31

A generalized Schrödinger formalism as a Hilbert space representation of a generalized Liouville equation

R K VARMA
Physical Research Laboratory, Ahmedabad 380 009, India

Abstract. A generalized Schrödinger formalism has been presented which is obtained as a Hilbert space representation of a Liouville equation generalized to include the action as a dynamical variable, in addition to the positions and the momenta. This formalism applied to a classical mechanical system had been shown to yield a similar set of Schrödinger like equations for the classical dynamical system of charged particles in a magnetic field. The novel quantum-like predictions for this classical mechanical system have been experimentally demonstrated and the results are presented.

Keywords. Classical and quantum mechanics; Hilbert space representation charged particle dynamics; non-Planckian discrete states.

PACS No. 03.65

1. Introduction

Quantum mechanics has, since its advent, revealed itself to be as enigmatic as it is novel. While it is a more general theory, supposed to be containing classical mechanics as a limit, it is not just a generalization of the latter, but represents a fundamentally distinct theory and indeed a paradigm shift from the conceptual framework of classical mechanics.

Attempts have continued since the inception of QM to forge a conceptual unity and continuity between the two theories through a possible unifying covering formalism. The present author has been concerned with the development of such a formalism for a number of years. Clearly, the most distinctive feature of quantum mechanics is the probability amplitude nature of its description and the interference effects resulting therefrom.

While there is obviously a genetic relationship between classical and quantum mechanics, there exist differences in their mathematical form and content which makes them appear so distinct from each other. Classical mechanics, for example, admits the whole continuum of initial values and the resulting dynamical states as 'allowed states', which are propagated as allowed states by the equation of motion. Quantum mechanics, by contrast, being a boundary value theory, permits, for a bounded system, only a discrete set of states of motion determined by the boundary conditions. Quantum mechanical states are thus determined by a set of global constraints (boundary conditions), while there are no such constraints for a classical mechanical system: A classical particle is not affected by a distant boundary, and represents only a local

evolution. One may, therefore, ask the question whether classical and quantum mechanics may be regarded as two different manifestations of the same underlying dynamics, described by the two different mathematical representations: global vs local.

There are two crucial facts about QM that need to be recalled. One, that it bears a close relationship with the Hamilton–Jacobi formalism of classical mechanics. The second one is that it is a probabilistic theory and ought to have its origin in a probabilistic framework. In classical statistical mechanics, the probability arises, from a lack of knowledge about the initial data. While the orthodox (Copenhagen) point of view ordains that quantum mechanics is intrinsically probabilistic, it has also been argued that the probability may be considered to arise because of a lack of knowledge about certain 'hidden variables' which when specified could make an event deterministic. We shall make two stipulations. One, that the probability in QM has the same origin as in classical statistical mechanics. Second, that the quantum events can be described by a classical Liouville equation, but in a suitably generalized form. To see the manner in which it is generalized we consider here in § 2 the Hamilton–Jacobi formalism via the theory of first order partial differential equation. In § 3, we give the generalized Liouville equation and in § 4 a generalized Schrödinger theory is obtained as a Hilbert space representation of the former. It will be seen that this formalism covers both classical and quantum mechanics suitably.

A fascinating consequence of this formalism, namely, obtaining Schrödinger theory as a Hilbert space representation of the classical (generalized) equation, is that this may be applied to a *suitably defined* classical mechanical system as well. This means that the particular classical mechanical system would be predicted to exhibit properties characteristic of a probability amplitude description. This would appear to be quite heretical. The interesting fact, however, is that such a classical mechanical system had indeed been identified by the present author more than two decades ago and a set of Schrödinger-like equations were obtained then (through a heuristic derivation) [1]. Later they were derived as a Hilbert space representation of the Liouville equation for the system [2]. Its quantum-like observable consequences have indeed been observed over the years. Some of these observed effects will be presented in § 4. In some sense these results are more fascinating than the derivation of the quantum mechanic Schrödinger equation, because they are something entirely new and unexpected and, in fact, cannot be understood in the framework of the standard classical mechanical paradigm. If the latter is indeed the case, as it seems to be, then we have an entirely new physics in the classical mechanical domain of parameters: What it means is that the standard classical mechanical paradigm need to be supplemented by constraints delimiting the allowed values of the initial conditions. It may be emphasized here that the predictions and subsequent observations of a set of entirely new phenomena which cannot be understood even *post-facto* in terms of the standard classical mechanical paradigm in whose domain they fall, do testify to the physical significance and reality of the Schrödinger-like description which led to these predictions. This may also serve as a validation of the procedure for obtaining the quantum mechanic Schrödinger formalism on a Hilbert space representation of a generalized Liouville equation which has been presented here and elsewhere earlier [3, 4].

2. The Hamilton–Jacobi formalism and a generalized Liouville equation in an extended phase space

The standard manner of arriving at the Hamilton–Jacobi formalism given in most standard text books on classical mechanics is via the canonical transformation. A generating function S which makes the new canonically transformed Hamiltonian vanish is essentially the Hamilton principal function and the defining equation for the new Hamiltonian is the Hamilton–Jacobi equation.

Following Courant and Hilbert [5] (for example) we introduce the Hamilton–Jacobi equation as a first order partial differential equation, whose characteristics are the Hamilton equation of motion. Thus, consider a partial differential equation of the form

$$\begin{aligned} &F(X_1, X_2, X_N, t, S; p_1, p_2, \ldots, p_N, p_t) \\ &\equiv p_t + H(X_1, X_2, \ldots, X_N, t; p_1, p_2, \ldots, p_N) = 0, \end{aligned} \tag{1}$$

where

$$p_i = \frac{\partial S}{\partial X_i}, \quad p_t = \frac{\partial S}{\partial t}$$

and where to correspond to the H–J equation, F has been taken to be independent of S, and the dependence on p_t to be linear.

With the above first order partial differential equation $F = 0$, one associates the following system of ordinary differential equations for the $(2N + 1)$ functions $(X_i, p_i = \partial S/\partial X_i, S)$ of a parameter τ,

$$\frac{\mathrm{d}X_i}{\mathrm{d}\tau} = \frac{\partial F}{\partial p_i} = \frac{\partial H}{\partial p_i}, \quad \frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{\partial F}{\partial p_t} = 1, \tag{2a}$$

$$\frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial F}{\partial X_i} - p_i \frac{\partial F}{\partial S} = -\frac{\partial H}{\partial X_i}, \tag{2b}$$

$$\frac{\mathrm{d}S}{\mathrm{d}\tau} = \sum_i p_i \frac{\partial F}{\partial p_i} + p_t \frac{\partial F}{\partial p_t} = \sum p_i \frac{\partial H}{\partial p_i} - H = L, \tag{2c}$$

where L is the Lagrangian. This system is referred to as the system of 'characteristic differential equations' belonging to the partial differential equation (1). Using the equation $\mathrm{d}t/\mathrm{d}\tau = 1$, the two sets of these equations (2a) and (2b) will be recognized as the Hamilton equations with H as the Hamiltonian function, while (2c) gives the time rate of change of the action S as the Lagrangian on the right. The partial differential equation (the H–J equation) and characteristic equations are thus equivalent to each other. It is this equivalence which was exploited by Jacobi to formulate his theory.

There is an advantage in the above manner of introduction of the H–J equation. It leads naturally to a complete set of characteristic equations: In addition to the standard Hamilton equations (2a) and (2b) one also finds eq. (2c), as one of the characteristic equations. This provides a basis for augmenting the phase space of a dynamical system, and writing down a generalized Liouville equation. As we shall see later, a Hilbert space representation of this generalized Liouville equation yields a generalized Schrödinger theory.

The set of 'characteristic differential equations' (2a,b,c) for the quantities (X_i, p_i, S) define trajectories in the space of the $(2N+1)$ variables. One can then construct a $(2N+1)$ dimensional phase space corresponding to these variables and a probability (phase) function $f(X_i, p_i, S)$ defined over this phase space. (Usually one defines a phase space of the $2N$ variables $(X_i p_i)$. But we include here also the variable S to correspond to the $(2N+1)$ characteristic equations (2a,b,c).) The 'equation of continuity' for $f(X_i, p_i, S)$ can then be written down as

$$\frac{\partial f}{\partial t} + \sum_i \frac{\partial}{\partial X_i}(\dot{X}_i f) + \sum_i \frac{\partial}{\partial p_i}(\dot{p}_i f) + \frac{\partial}{\partial S}(\dot{S} f) = 0. \tag{3}$$

Using the expressions (2a) and (2b) for $\dot{X}_i$ and $\dot{p}_i$ and the fact that $\dot{S} = L$ is independent of S, so that $\partial \dot{S}/\partial S = 0$, eq. (3) becomes

$$\frac{\partial f}{\partial t} + \sum \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial X_i} + \sum \frac{\partial H}{\partial X_i}\frac{\partial f}{\partial p_i} + L\frac{\partial f}{\partial S} = 0. \tag{4}$$

This is a generalization of the classical Liouville equation, in the extended phase space of (X_i, p_i, S) resulting in the inclusion of the last term, $L\partial f/\partial S$. This, as usual, describes the conservation of the probability measure along the characteristics described by (2a,b,c). The underlying dynamics is, of course, the classical Hamiltonian flow. Since all the 'flow' components $(\partial H/\partial p_i, -\partial H/\partial X_i, L)$ are independent of S, an integration with respect to S yields the standard classical Liouville equation

$$\frac{\partial \bar{f}}{\partial t} + \frac{\partial H}{\partial p_i}\frac{\partial \bar{f}}{\partial X_i} - \frac{\partial H}{\partial X_i}\frac{\partial \bar{f}}{\partial p_i} = 0 \tag{5}$$

with

$$\bar{f} = \int \mathrm{d}S\, f. \tag{6}$$

2.1 *Classical analogue of the state of motion in quantum mechanics*

Dirac, as the readers would already know, was quite occupied with the question of the nature of relationship between classical and quantum mechanics, and with the question of classical analogues of quantum objects. In his 1933 paper [6], for example, he very beautifully elucidated the role of classical Lagrangian in quantum mechanics, which later culminated into the path integral formulation at the hands of Feynman [7]. Another important remark, made by Dirac [8] was in relation to the significance of the Hamilton principal function as a solution of the Hamilton–Jacobi equation. A principal function, satisfying the Hamilton–Jacobi equation, defines a 'family' [11]. "The family does not have any significance from the point of view of Newtonian mechanics; but it is a family which corresponds to one state of motion in the quantum theory, so presumably the family has some deep significance in nature, not yet properly understood" (quoted from Dirac [8]).

Coming to the generalized Liouville equation (4) the Liouville density function f can be any function of X_i, p_i and S, which solves (4). In my earlier papers, however, f was chosen to be a δ-function in all the initial momenta including the global integrals of

motion. This was done to make the distribution f correspond closely to a quantum state where the momenta (in terms of which the distribution is a δ-function) represent integrals of motion with respect to which the state is prepared. I was not aware of Dirac's remarks then, but it turns out that such a distribution represents what Dirac has called a 'family', defined by a principal function, and which according to him corresponds to one state of motion in the quantum theory.

In accordance with the above considerations, f is regarded as a function of the initial momenta values α_i, rather than of the current momenta p_i. Equation (4) is then transformed accordingly (from the current momenta p_i to the initial momenta α_i). This yields

$$\frac{\partial \hat{f}}{\partial t} + v_i \frac{\partial \hat{f}}{\partial X_i} + \frac{L}{\eta} \frac{\partial \hat{f}}{\partial \Phi} = 0, \tag{7}$$

where now $\hat{f} = \hat{f}(X_i, \Phi, t; \alpha_i)$ and v_i is not to be regarded as independent of the X_i, but as a function of the X_i and the α_i, while the α_i appear as parameters in the argument of $\hat{f}$. Here, we have also introduced a constant η of the dimensions of action to non-dimensionalize the action S, so that

$$S = \eta \Phi \tag{8}$$

or to measure S in units of η. Φ is then a dimensionless quantity which may be termed as 'action-phase'. The meaning of $\hat{f}$ is then also analogous to the meaning of probability in quantum mechanics: $\hat{f}$ gives the probability of finding the system at X_i at time t with a given value of S, if it had initially the momenta α_i.

2.2 *A Hilbert-space representation*

Equation (7) with the definition (8) and the underlying characteristic equations (2a,b,c), it may be recalled, are essentially the equations representing classical dynamics, though the Liouville equation has been generalized to include a term corresponding to the action variable S. It may be recalled that eq. (5) for $\hat{f}$, the action averaged distribution is essentially the classical Liouville equation. One seeks a Hilbert space representation of eq. (7) and writes accordingly:

$$\hat{f} = \psi^2, \tag{9}$$

where ψ is a real quantity (one could let ψ be complex and write $\hat{f} = \psi^* \psi$, but there is no loss of generality if ψ is taken to be real). Next a finite time representation of eq. (7) is written as

$$\hat{f}(X_i, \Phi, t; \alpha_i) = \hat{f}\Bigg[X_i - \int_{t'}^{t} v_i(X_i(t''), \alpha_i) \mathrm{d}t'' \\ \Phi - \frac{1}{\eta} \int_{t'}^{t} L(X_i(t''), v_i(t'')), \mathrm{d}t'', t', \alpha_i\Bigg]. \tag{10}$$

Using (9) and taking the square root of (10), choosing the position sign at all space-time points:

$$\psi(X_i, \Phi, t; \alpha_i) = \psi\left(X_i - \int_{t'}^{t} v_i(t'') \mathrm{d}t'', \Phi - \frac{1}{\eta} \int_{t'}^{t} L \, \mathrm{d}t'', t', \alpha_i\right). \tag{11}$$

I shall not repeat the procedure given elsewhere [3, 4] which yields, starting from (11), a generalized set of Schrödinger equations

$$\frac{i\hbar}{n}\frac{\partial\Psi(n)}{\partial t} = -\left(\frac{\hbar}{n}\right)^2 \frac{1}{2m}\nabla^2\Psi(n) + V\Psi(n), \quad n = 1, 2, 3, \ldots \tag{12}$$

with the total probability density

$$G(X, t) = \sum_n \Psi^*(n)\Psi(n). \tag{13}$$

It is, however, pertinent to point out some crucial steps, which represent a point of departure from classical mechanics and which lead to the above set of Schrödinger equations.

The first is clearly the generalization of the Liouville equation to include the action S as a variable. However, as the action-averaged Liouville equation (5) (which is the standard classical Liouville equation) shows, the classical mechanics is clearly contained in (4).

The second step is the eq. (9) writing $\hat{f} = \psi^2$ seeking a Hilbert space representation of eq. (7). While it may be an unusual proposition to seek a Hilbert space representation of (7), clearly this by itself could not make it depart in its content from that of classical mechanics, for one could in principle solve for ψ and reconstruct $\hat{f} = \psi^2$ and action-average it to obtain the classical distribution $\bar{f}$.

The third step is the assumption of ψ being periodic in $\Phi = S/\eta$ with the period 2π. It is this assumption (where η is identified with $\hbar$) which enable ψ to be Fourier-decomposed as a series (rather than an integral) and the subsequent procedure to be followed leading up to the derivation of the Schrödinger equations which are second order in space. One should follow the derivation closely to see how the second order operator ∇^2 arises. It is clear that this operator could not have arisen if the action S were not included as a variable.

The important consequence of this structure (second order differential operator) is that the functions $\Psi(n)$ of eq. (12) should be specified through appropriate boundary conditions. This amounts, apart from other consequences, to delimiting the allowed values of certain integrals of motion to a discrete set – that is, to 'quantization'.

It ought to be pointed out that this particular derivation yields a generalized formalism in the form of an infinite set of Schrödinger equations [12] with $(\hbar/n)$, $n = 1, 2, 3 \ldots$ in the place of $\hbar$ and a correspondingly generalized probability connection (13). Some of the direct observable consequences of these additional modes $n = 2, 3, \ldots$ have been discussed in ref. [4]. They have not been so far identified nor has there been any attempt made to check their existence experimentally.

3. A classical mechanical system with the Schrödinger-like description

There is nothing special in the procedure given above (where we obtained the Schrödinger-like description as a Hilbert space representation of the Liouville equation) which could not be applied to any suitable classical mechanical system. As pointed out above perhaps the most important step where the quantization really came in was

the assumption of periodicity of ψ with respect to $\Phi(\equiv S/\hbar)$ with the period 2π. What is, of course, required for the classical mechanical system is the justification of periodicity with respect to the action phase and the identification of an appropriate unit of action (analogous to $\hbar$) belonging to the system in terms of which to define the latter.

Such a classical mechanical system was identified by the present author more than two decades back [1]. The dynamical system of charged particles in a magnetic field is such a system. The action in the system which can serve as a unit is the gyroaction $\mu = \frac{1}{2} m v_\perp^2/\Omega$, where $v_\perp$ is the component of velocity perpendicular to the magnetic field at the point of injection and $\Omega = eB/mc$ is the gyro-frequency also at the point of injection, so that μ is essentially an initial value which is an exact constant of motion.

The Liouville equation for the system is given by

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial X} - \Omega \frac{\partial f}{\partial \phi} = 0, \tag{14}$$

where ϕ is the gyro-phase.

Transforming the equation to momenta initial values: the total energy E, the canonical angular momentum P_ϕ, and the gyro-action μ, as well as transforming ϕ to the action phase

$$\begin{aligned} \Phi &= \frac{1}{\mu} \int_0^t \frac{1}{2} m v_\parallel^2 \mathrm{d}t' + \phi \\ &= \frac{1}{\mu} \int_0^t \mathrm{d}t' \left(\frac{1}{2} m v_\parallel^2 - \mu\Omega \right) \\ &= S_A/\mu, \end{aligned} \tag{15}$$

where S_A is the time integral of the adiabatic Lagrangian $L_A = (\frac{1}{2} m v_\parallel^2 - \mu\Omega)$, one obtains

$$\frac{\partial f}{\partial t} + v_\parallel \frac{\partial f}{\partial X_\parallel} + \frac{L}{\mu} \frac{\partial f}{\partial \Phi} = 0. \tag{16}$$

This equation has a structure similar to that of eq. (7) except that μ takes the place of η (later identified as $\hbar$). A Hilbert space representation of this equation which can be sought without any prejudice to QM, yields in a similar manner, the following set of Schrödinger-like equations:

$$\frac{i\mu}{n} \frac{\partial \hat{\Psi}(n)}{\partial t} = -\frac{(\mu/n)^2}{2m} \frac{\partial^2 \hat{\Psi}(n)}{\partial X_\parallel^2} + (\mu\Omega)\hat{\Psi}(n), n = 1, 2, 3, \ldots \tag{17}$$

again with the total probability density

$$G(X_\parallel, t) = \sum_n \hat{\Psi}^*(n)\hat{\Psi}(n). \tag{18}$$

The periodicity of $\psi = \sqrt{f}$ with respect to Φ which is crucial for obtaining the Schrödinger-like equation (17) is guaranteed through the periodicity with respect to ϕ (the gyro-phase) which is an additive part of Φ as per eq. (15).

It may be mentioned that $(\mu\Omega)$ which appears in the place of 'potential' is in fact the adiabatic potential which describes the effective parallel motion along the magnetic field through the equation of motion (in the adiabatic approximation)

$$m \frac{\mathrm{d}v_{\parallel}}{\mathrm{d}t} = -\nabla_{\parallel}(\mu\Omega). \tag{19}$$

4. Quantum-like consequences for the classical mechanical system

Even though the above mentioned dynamical system of charged particles in a magnetic field is a classical mechanical one with macroscopic dimension, the above amplitude description governed by the Schrödinger-like equations (18) entails that this system exhibits quantum-like properties for the ensemble of trajectories (the coherent set of trajectories, á la Synge [9] or a 'family' á la Dirac [8]), for which these equations are obtained. Whether this is indeed the case can be ascertained only through appropriate experimentation. Two kinds of predictions were identified for experimental verification.

4.1 *Existence of multiplicity of residence times in an adiabatic trap*

One relates to the residence times of charged particles trapped in an adiabatic potential well as described by the adiabatic equation of motion (19). Charged particles can be confined in a well of the potential $\mu\Omega$, if the total energy E of the particle is $E < (\mu\Omega)_{\max}$. But the trapping is not for ever, because the equation of motion (19) which describes this trapping is only an approximate equation. Particles so trapped have indeed been found experimentally to have leaked out of the trap, with characteristic residence times. No satisfactory theory for the determination of these residence life-times existed when these experimental results were reported during 1968–1969. The Schrödinger-like equations (17) along with (18), which were first given by the author in 1971 as a possible description of the leakage of particles, afforded the possibility of determining the life times through a quantum-tunneling like process. Leaving aside for the moment the question of the interpretation of the amplitude functions $\hat{\Psi}(n)$ vis-a-vis the quantum mechanics Schrödinger wave functions, what is different here from the standard quantum formalism is that we have here not one but a set of equations for the $\hat{\Psi}(n)$ with $n = 1, 2, 3....$ This, as already pointed out, is also true of the generalized Schrödinger formalism for quantum mechanics as contained in the set of equations (12) and (13). Correspondingly, one predicts here the existence of a multiplicity of residence times as determined by the equations corresponding to $n = 1, 2, 3 \ldots$, for the injection of particles with the same energy and the pitch angle. It should, perhaps, be pointed out that no predictions of this kind were available based on the standard approach.

Such multiplicity of residence times were indeed found experimentally in a series of experiments carried out at the Physical Research Laboratory [10a, b, c]. Upto three distinct times were identified, and were found to correspond to the modes $n = 1, 2, 3$ in their dependence on the magnetic field strength for a given energy E and the pitch angle

of injection θ. The experimental identification of these modes $n = 1, 2, 3$ through the existence of the corresponding residence times then constitutes a validation of the Schrödinger-like description. It may also be mentioned that so far no way has been found in terms of the standard approach to the problem to describe these multiplicities of residence times.

4.2 *Existence of discrete allowed and forbidden states in the macro-domain*

The other kind of prediction refers to the wave-like manifestations of the charged particle dynamics in the macro domain, that is the dimensions of tens of centimeters. Such manifestations follow from the wave nature of the Schrödinger-like equations, which include the existence of interference effects and consequent discrete allowed and forbidden energy states in the classical domain of parameters. Astonishing as it may seem such non-Planckian discrete energy states have indeed been observed in recent experiments, as will be described below.

These observations appear to be in manifest violation of the Lorentz equation of motion – initial value paradigm since the latter permits all the continuum of initial values and hence energies. As a matter of fact, the observations demonstrate in conformity with the predictions of the theory, a kind of wave-like nonlocality whereby the relative separation between the discrete states depends on the distance between the electron source – the electron gun, and the detector. Such a distance dependence is quite characteristic of the quantum behaviour, but is completely inexplicable in terms of the standard paradigm, whereby the entire continuum of initial values are allowed values and which takes no cognizance of the distant boundaries.

It seems therefore, that an electron moving in a magnetic field exhibits a wave-like behaviour even in macroscopic dimensions very much like the wave nature of electron in micro-dimensions (de Broglie wave). Such a behaviour follows from the set of Schrödinger-like equations obtained for the system.

What is then the resolution of the apparent paradox? The Lorentz equation is certainly not wrong. But the experiments, on the other hand, cannot be explained in terms of the standard paradigm, but appear to conform to the Schrödinger-like description.

The resolution of the parodox appears to consist in the realization that the standard paradigm of classical mechanics represents a purely 'local' evolution. That is, the trajectory evolution at a given time t is determined entirely by the value of the field of force at the position of the particle at that time, that is 'locally'. The Schrödinger-like equation on the other hand, seems to provide a global description of the same dynamics. Such a global description takes cognizance of the distant boundaries. As a consequence certain energy values (as initial conditions) are found to be disallowed, as deduced from the theory and verified experimentally. The allowed values on the other hand, could well propagate according to the Lorentz equation of motion. If this point of view is taken there need not be any violation of the Lorentz equation, but only of the paradigm that all the continuum of initial values are the allowed values in the classical mechanical formalism.

It may not be possible for reasons of space to present here the derivation of the formula giving the discrete allowed states in terms of a 'quantum number'. If we consider only the function $\Psi(1)$ for $n = 1$, corresponding to an electron beam from an electron gun

propagating along a magnetic field, then it has been shown (Varma [11]) that the probability density of finding the electrons on a grounded detector after propagating a distance D along the field is given by

$$\Psi^*\Psi = \sum_k \frac{|\hat{\psi}(k)|^2}{k} + \sum_k 2R(K,\bar{k}) \sin\left[\frac{\bar{\Omega}D}{\bar{v}_\parallel} + \phi\right], \tag{20}$$

where $v_\parallel$ is the 'parallel' velocity

$$v_\parallel = [2(E - \mu\Omega)/m]^{1/2}. \tag{21}$$

We see that this expression has an oscillating term (the second term) (over the mean first term) which oscillates with the magnetic field B, with the energy E, and the distance traversed D individually with the other parameters remaining constant. The oscillating term is a consequence of the interference effects. The minima in the expression of $\Psi^*\Psi$ are interpreted as 'forbidden states' of the charged particles in a magnetic field. The energies of the 'forbidden states' are given by

$$E_j = \tfrac{1}{2}m(\Omega D/2\pi)^2/(j + \phi_j/2\pi)^2, \tag{22}$$

where j denotes the 'quantum number' labelling the forbidden energy states E_j with ϕ_j being the phase shift.

As is clear from the expression (22) for the 'forbidden energies', they are seen to be determined, very astonishingly, by the distance D over which the electrons travel from the gun to the detector. This is quite astonishing from the point of view of the standard paradigm of classical mechanics. For, the dependence on D signifies a kind of wave-like nonlocality while classical mechanics is known to be a local theory. The predictions contained in this relation were subjected to experimental scrutiny at the Physical Research Laboratoy. We describe below an experiment reported earlier [12(a, b)] to check these predictions.

5. The experiment and its results

An electron beam of very low intensity ($< 0.1\,\mu$A) is injected from an electron gun along a magnetic field ($\sim$200–300 gauss) in an SS vacuum chamber of 27 cm diameter evacuated to $\sim 5.10^{-7}$ torr, and is received at a detector (Faraday cup) at a distance ($L \sim 20$ cm) away from the gun. The injection is almost parallel to the field. The Faraday cup detector consists of a grounded collector plate of $\sim$ 25 cm diameter, 1 cm behind a grid which can be biased to any required potential. The electron current from the beam can be measured by the detector (current received by the plate) for various negative grid potentials from zero to $-\Phi_{\max}$, where $\Phi_{\max} > E/|e|$, E being the electron beam energy. Since according to classical mechanics, the whole continuum of energy states are allowed states, the plate current is expected to exhibit a monotonically increasing response as the negative grid potential is swept from $-\Phi_{\max}$ to zero.

The actual experimentally observed response is quite and astonishingly different from the expected one (à la the standard classical mechanical paradigm) and is shown in

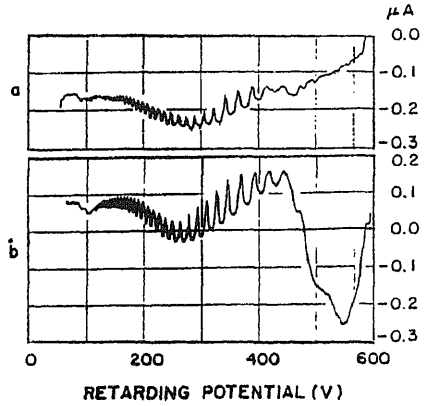


Figure 1. Plate current (**a**) and grid current (**b**) as functions of the retarding potential. $\bar{B} = 170\,\text{G}$, $L = 30\,\text{cm}$ and $E = 600\,\text{eV}$.

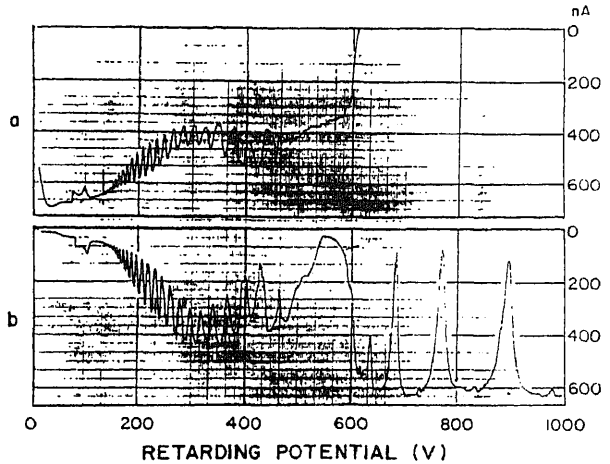


Figure 2. Plate current (**a**) and anode current (**b**) as functions of the retarding potential $\bar{B} = 177\,\text{G}$, $L = 19\,\text{cm}$ and $E = 600\,\text{eV}$.

figure 1. The upper curve (a) is the plate current and quite clearly exhibits a series of sharply defined dips. The lower curve (b) is the grid current for the same set of parameters and also exhibits a series of dips which are found to be exactly correlated with the plate current dips. From the point of view of classical mechanics, these dips are quite unexpected and astonishing as they signify the existence of 'forbidden states' which have no place in the formalism of classical mechanics.

To further check that these are indeed the forbidden states, both the plate current and the anode current were measured simultaneously as the grid potential is swept. (It should be explained that the anode which is grounded, is a part of the electron gun and which accelerates the electrons emanating from a negatively biased hot cathode placed about one cm away from it.) Thus the electrons encounter the anode much before ($\sim 20\,\text{cm}$) on the way to grid-plate assembly. Figure 2 shows both the simultaneously recorded plate

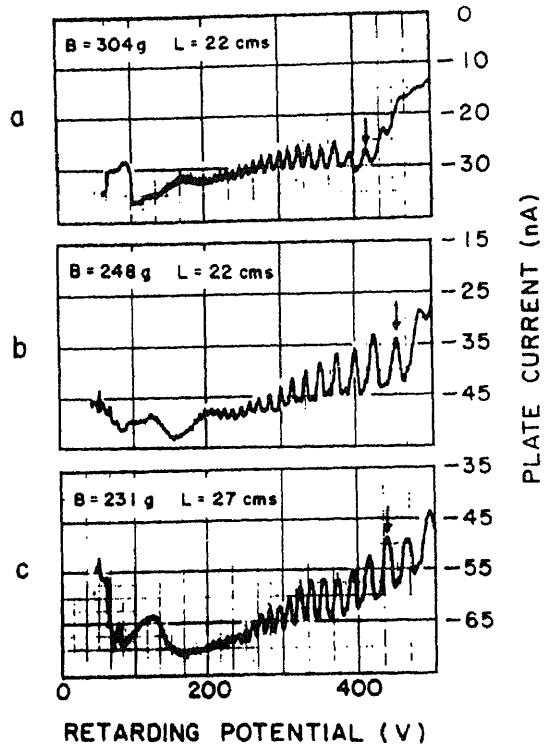


Figure 3. Plate current as a function of the retarding potential for different values of $\bar{B}$ and L (as shown in the plots) and $E = 650$ eV.

Table 1. The energies E_j and the j values for plots in figure 3(a, b, c).

Peaks	Energy E_j (eV)			$j + \phi/2\pi$		
	Plot 3a	Plot 3b	Plot 3c	Plot 3a	Plot 3b	Plot 3c
$N(\downarrow)$	417	453	438	$41 + 0.15$	$31 + 0.32$	$36 + 0.56$
$N+3$	357	377	373	$44 + 0.48$	$34 + 0.36$	$39 + 0.62$
$N+6$	313	317	323	$47 + 0.45$	$37 + 0.47$	$42 + 0.57$
$N+9$	277	272	283	$50 + 0.50$	$40 + 0.46$	$45 + 0.47$
$N+12$	247	237	250	$53 + 0.48$	$43 + 0.35$	$48 + 0.41$

and anode currents, the upper curve being the plate and the lower one being the anode current. The two curves are clearly anti-correlated, showing that electrons that did not (or could not) pass through the system in a certain energy state (forbidden state) have found their way to the anode, confirming in a sense the interpretation of the 'dips' in the plate current as forbidden states.

The experiment was repeated with different distances between the gun and the detector and magnetic fields. The plots so obtained are shown in figure 3. We see that the positions of the dips change both with the distance L and the magnetic field B. A little reflection

shows that the dependence on the distance L is most enigmatic because it signifies a kind of wave-like nonlocality.

One may now finally examine whether the dips so observed can be described by the relation (23). To do so, we read the positions of the dips (in energy) from the plots and using these values on the left of eq. (23) with the magnetic field and distance D inserted on the right, we calculate the quantity $(j + \phi/2\pi)$ for every third dip counted from an (arbitrary) arrowed peak in three respective plots of figure 3. The results are shown in table 1. (It should be pointed out that $D = 3L$ was used in the calculation for reasons to be explained elsewhere.) The whole number in the value of $(j + \phi/2\pi)$ so obtained is identified with the 'quantum number' characterizing the dip and the fraction with $(\phi/2\pi)$. It is clearly seen that the j values do differ by 3 corresponding to the fact that every third dip was chosen. Since the different curves correspond to different B and L values, the dependence on B and, in particular, on the distance L, is well borne out by the experimental results.

In summary, it is important to highlight the following facts that we have demonstrated:

a) The discrete forbidden states of motion do exist in the domain of parameters where one would use classical equation of motion to determine the motion.
b) The energies of these 'forbidden states' are well represented by the relation (5) which is obviously nonquantal as there is no Planck quantum h appearing in it.
c) The forbidden states E_j form a hydrogen-like sequence for which 'quantum numbers' j and the phases can be identified as in table 1.
d) The forbidden states E_j and the associated quantum numbers j depend on the distance between the gun and the detector. This is a manifestation of wave-like behaviour which is not known to be a characteristic of the standard initial value paradigm of classical mechanics.

Based on the above described experimental results, one may now conclude that the elèctrons moving in a magnetic field do appear to exhibit a wave-like behaviour in macroscopic dimension which is known to be a domain of operation of classical dynamics and which, therefore, admits a continuum of energy states. The existence of discrete forbidden states as described above appears to be in contradiction with the latter.

6. Implication of these results

To summarize, it has been predicted theoretically and demonstrated experimentally, that electrons moving along a magnetic field within the classical mechanical domain of parameters, exhibit the existence of discrete forbidden states, a feature which is not admitted by the standard initial value paradigm of classical dynamics. The latter admits the entire continuum of initial conditions as allowed states. This is an entirely new feature of classical dynamics, not hitherto pointed out to the best of the author's knowledge. What is most astonishing and violative of the classical mechanical intuition is that the forbidden (as also the allowed) states are determined nonlocally by the distant boundaries. The allowed states may well evolve à la the Lorentz equation of motion. The new description à la the Schrödinger-like equation merely constrains the allowed initial data.

The Schrödinger-like formalism may thus be regarded as providing a global description of the classical mechanical system as it takes cognizance of the boundaries which the standard paradigm does not. It is of course, interesting and significant that a classical mechanical system, such as the one under consideration, admits a Schrödinger wave-like description, so far considered as the sole preserve of quantum mechanics. Of course, not every classical mechanical system may be so describable. We have not been able to identify general criteria for such systems.

We have thus shown how the generalized Liouville equation for the coherent system gives the Schrödinger formalism (generalized) as its Hilbert space representation provided the Liouville density and the square-root thereof is periodic in the action phase $\Phi = S/\hbar$. The classical Liouville equation is obtained as an action-phase average of the generalized Liouville equation. It is in this manner that classical mechanics is contained in this generalized formalism. It will be seen that formally the limit $\hbar \to 0$ (which is usually identified as the classical limit) also amounts, in the lowest order, to an averaging over Φ.

The question that is most likely to be raised is that of the validity and the validation of the formalism obtained and the essential steps and arguments leading to it. Such a validation of the formalism has been provided in a rather spectacular fashion.

As discussed earlier the main thrust of our investigation was to provide a formalism which could cover both classical and quantum mechanics in a more natural and logically connected manner. The same formalism applied to an appropriate classical mechanical system (charged particle dynamics in a magnetic field) has yielded a Schrödinger-like description whose quantum-like predictions have been experimentally verified in a rather spectacular fashion. These are astonishing observations in as much as they cannot be understood in terms of the standard classical mechanical paradigm. Our line of investigation has thus helped unearth entirely new physical effects in the classical mechanical domain of parameters. More than just that, they seem to point to a new paradigm in classical mechanics whereby not the whole of the continuum of initial values are allowed values, just like in the quantum domain.

The question that presents itself is that what is it that leads a classical mechanical system, to the discreteness of the allowed states as against the continuum as per the standard paradigm. It has been shown by the author [13] recently that the topological considerations of the system configuration space leads to conditions which constrain the allowed initial values to a discrete set. These conditions turn out to be nothing but the eigenvalue equations in the EBK form of the Schrödinger-like equations (17). It is conjectured that the same that is, the topology, may also be the origin of discreteness of allowed energy values in quantum mechanics.

These investigations thus provide another new point of view in terms of which to consider the question of classical-quantum relationship: Quantum mechanics as a Hilbert space representation of the (generalized) classical Liouville equation may thus be considered as a global description of the same dynamics, with the boundary conditions simply helping to determine the wave functions and corresponding discrete energies as eigenvalues permitted by the overall topology of the system. This in turn translates into an appropriate Liouville density function being determined by the boundary conditions.

References

[1] R K Varma, *Phys. Rev. Lett.* **26**, 917 (1971)
[2] R K Varma, *Phys. Rev.* **A31**, 3951 (1985)
[3] R K Varma, *Had. J. Suppl.* **1**, 437 (1985)
[4] R K Varma, in *Microphysical reality and quantum formalism* edited by A van der Merve (Kluwer Academic Publishers, 1988) p. 175
[5] R Courant and D Hilbert, *Methods of mathematical physics* vol. II, p. 107 (Interscience Publishers, 1962)
[6] P A M Dirac, *Z. Phys. Sov. Union* **3**, 63 (1933)
[7] R P Feynman, *Rev. Mod. Phys.* **20**, 367 (1948)
[8] P A M Dirac, *Can. J. Math.* **3**, 1 (1951)
[9] J L Synge, in *Principles of mechanics and field theory*, Handbuch der Physik, vol. III/1, Springer Verlag (Berlin, Gottingen, Heidelberg) (1960), p. 121
[10] D Bora, P I John, Y C Saxena and R K Varma, (a) *Phys. Lett.* **A75**, 60 (1979); (b) *Plasma Phys.* **22**, 653 (1980); (c) *Phys. Fluids* **25**, 2284 (1982)
[11] R K Varma, in *A variety of plasmas*, Proc. Int. Conf. Plasma Phys., Delhi, Indian Acad. of Sciences (1989) p. 235
[12] (a) R K Varma and A M Punithavelu, *Mod. Phys. Lett.* **A8**, 169 (1993)
(b) R K Varma and A M Punithavelu, *Mod. Phys. Lett.* **A8**, 3823 (1993)
[13] R K Varma, *Mod. Phys. Lett.* **A9**, 3653 (1994)

—journal of
physics

July 1997
pp. 33–40

The quantum geometric phase as a transformation invariant

N MUKUNDA*
Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560 012, India
*Honorary Professor, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064, India

Abstract. The kinematic approach to the theory of the geometric phase is outlined. This phase is shown to be the simplest invariant under natural groups of transformations on curves in Hilbert space. The connection to the Bargmann invariant is brought out, and the case of group representations described.

Keywords. Geometric phase; Bargmann invariant.

PACS No. 03.65

1. Introduction

Ever since Berry's important work of 1984 [1], the geometric phase in quantum mechanics has been extensively studied by many authors. It was soon realised that there were notable precursors to this work, such as Rytov, Vladimirskii and Pancharatnam [2]. Considerable activity followed in various groups in India too, notably the Raman Research Institute, the Bose Institute, Hyderabad University, Delhi University, the Institute of Mathematical Sciences to name a few.

IIn the account to follow, a brief review of Berry's work and its extensions will be given [3]. We then turn to a description of a new approach which seems to succeed in reducing the geometric phase to its bare essentials, and which is currently being applied in various situations [4]. Its main characteristic is that one deals basically only with quantum kinematics. We define certain simple geometrical objects, or configurations of vectors, in the Hilbert space of quantum mechanics, and two natural groups of transformations acting on them. We then seek the simplest expressions invariant under these transformations. This is seen to lead immediately to a kinematic view of the geometric phase. It helps us pose useful questions and connect the phase to other more familiar expressions in quantum mechanics.

2. Quantum kinematics: States, phases and rays

As is well-known, in the quantum mechanical description of any system the dynamical variables are linear operators on a suitably constructed Hilbert space $\mathcal{H}$, while the various (pure) states of the system correspond to vectors $\psi, \psi', \ldots$ in $\mathcal{H}$. The superposition

principle of quantum mechanics permits us to form linear combinations $c_1\psi_1 + c_2\psi_2 \ldots$ of given vectors $\psi_1, \psi_2 \ldots$ using complex coefficients $c_1, c_2 \ldots$, and thus to produce new pure states from old ones in a completely nonclassical way. (This is of course apart from superselection rules.) The relation between vectors and physical states is quite subtle. Namely, a change in the overall phase of a vector in $\mathcal{H}$, $\psi \to e^{i\alpha}\psi$, causes no change in the physical state at all. Thus the relation between vectors in $\mathcal{H}$ and distinct physical states is many-to-one. One is thus motivated to introduce the concept of rays [5] – equivalence classes of vectors differing from one another by phases alone–so that between rays and physical states the relation is one-to-one. A given vector $\psi \in \mathcal{H}$ determines a corresponding ray consisting of all the vectors $e^{i\alpha}\psi$, $0 \leq \alpha < 2\pi$; this ray can be conveniently represented by the pure state density matrix or projection operator onto ψ, namely $\rho = \psi\,\psi^\dagger$, and one has these relationships:

$$\begin{aligned}
&\text{Vectors } \psi \text{ in } \mathcal{H} \overset{\text{many-to-one}}{\longrightarrow} \text{ physical states;} \\
&\text{Vectors } \psi \text{ in } \mathcal{H} \overset{\text{quotient w.r.t } U(1)}{\longrightarrow} \text{ rays } \rho = \psi\psi^\dagger; \\
&\text{rays } \rho = \psi\,\psi^\dagger \overset{\text{one-to-one}}{\longrightarrow} \text{ physical states.}
\end{aligned} \tag{1}$$

The geometry of ray space is somewhat complicated, and not as easy to visualize as that of $\mathcal{H}$. Moreover the superposition principle is not manifest, though it is certainly present. Nevertheless, its use is crucial not only in the above context but also for the geometric phase.

In contrast to overall phases being unobservable, relative phases in a linear combination, such as in $\psi_1 + e^{i\alpha}\psi_2$, are physically observable – all quantum mechanical interference phenomena ultimately have their origins in such relative phases.

We shall hereafter deal with unit vectors in $\mathcal{H}$, namely with points on the unit sphere in $\mathcal{H}$, and will not mention this repeatedly.

3. Quantum dynamics, adiabatic theorem and the geometric phase

Quantum dynamics is governed by the time-dependent Schrodinger equation: for each t we have a vector $\psi(t) \in \mathcal{H}$ evolving according to

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\psi(t) = H(t)\psi(t). \tag{2}$$

Here $H(t)$ is the Hamiltonian operator of the system, and the possible explicit time dependence allows for external or environmental influences on the system. For a completely isolated system there would be no such dependence.

Berry's original discovery of the geometric phase was in the following context. Imagine a situation where $H(t)$ changes very slowly – adiabatically – with passage of time, and suppose for each time we have a nondegenerate normalized eigenvector $u(t)$

with corresponding eigenvalue $E(t)$:

$$H(t)\, u(t) = E(t)u(t). \tag{3}$$

Thus we have an instantaneous energy eigenstate with time-dependent energy. Then the Born–Fock adiabatic theorem of quantum mechanics [6] states that an approximate solution of the Schrodinger equation (2) is given by

$$\psi(t) \simeq \exp\left[\frac{-i}{\hbar}\int_0^t E(t')\mathrm{d}t'\right]u(t), \tag{4}$$

where the (hitherto) free phase factor in $u(t)$ is restricted by the convention

$$(u(t),\, \dot{u}(t)) = 0. \tag{5}$$

Now suppose we have a cyclic environment: $H(T) = H(0)$ for some time T. The question is whether, when the environment returns to its original condition, the state vector also does so. While it was long known that the physical state returns to its original form, the precise behaviour of the state vector had not been properly appreciated until Berry's work. What Berry found was that there was a new piece in the phase factor relating $\psi(T)$ to $\psi(0)$, which was geometrical in origin:

$$\begin{aligned}
\psi(T) &\simeq \exp\left(\frac{-i}{\hbar}\varphi_{\mathrm{tot}}\right)\psi(0),\\
\varphi_{\mathrm{tot}} &= \varphi_{\mathrm{dyn}} + \varphi_{\mathrm{geom}},\\
\varphi_{\mathrm{dyn}} &= \int_0^T \mathrm{d}t\, E(t).
\end{aligned} \tag{6}$$

The total phase φ_{tot} is made up of a dynamical part, already evident in eq. (4) expressing the adiabatic theorem, and a geometric piece φ_{geom} which cannot be transformed away. Indeed, one can see that a nontrivial geometric phase implies $u(T) \neq u(0)$ if eq. (5) is obeyed.

Thus the geometric phase was originally found in the context of adiabatic, cyclic unitary evolution governed by the Schrödinger equation. Later developments showed that each of the first three conditions could be relaxed [7] – thus the geometric phase could be usefully defined even for nonadiabatic, noncyclic, nonunitary evolution governed by the Schrödinger equation. The developments to be now described show that one can even dispense with the Schrödinger equation!

4. The kinematic approach

On the unit sphere in $\mathcal{H}$ we define the following geometric objects: Open smooth parametrized curves C, which may be pictured as strings lying on the sphere. Such a curve C may be represented as a collection of vectors in this way:

$$C = \{\psi(s),\ s_1 \leq s \leq s_2\}. \tag{7}$$

The end points are $\psi(s_1)$ and $\psi(s_2)$; and while $\psi(s)$ must vary continuously with no breaks, a finite number of points where it is not differentiable can be tolerated.

Two groups of transformations can be defined to act on these curves, mapping each $\mathcal{C}$ to another $\mathcal{C}'$:

(i) Local phase changes:

$$\mathcal{C} \to \mathcal{C}' : \psi(s) \to \psi'(s) = \mathrm{e}^{i\alpha(s)}\psi(s), \qquad \alpha(s) = \text{smooth function of } s. \tag{8}$$

(ii) Monotonic reparametrizations:

$$s' = f(s), \quad \dot{f}(s) > 0, \qquad \psi'(s') = \psi(s). \tag{9}$$

If we denote the ray space image of $\mathcal{C}$ by C, it is clear that the local phase changes do not alter this image at all. The reparametrization transformation traverses the same set of points but at an altered rate; so the set of points comprising $\mathcal{C}$ receives a new description. The same is then true of C as well.

We now ask for the simplest functional of $\mathcal{C}$ with the property that it is invariant under both these transformations. With a little bit of work, one easily finds that the following expression is doubly invariant [8]:

$$\varphi_g = \arg(\psi(s_1), \psi(s_2)) - \int_{s_1}^{s_2} \mathrm{d}s\, \mathrm{Im}(\psi(s), \dot{\psi}(s)). \tag{10}$$

This is the geometric phase associated with $\mathcal{C}$! We see that it is the difference of two terms, each individually dependent on $\mathcal{C}$. However local phase invariance means that φ_g is really a functional of the ray space image C of $\mathcal{C}$; and the reparametrization invariance indicates its geometric nature. All this can be expressed as follows:

$$\begin{aligned} \varphi_g[C] &= \varphi_{\mathrm{p}}[C] - \varphi_{\mathrm{dyn}}[C], \\ \varphi_{\mathrm{p}}[C] &= \arg(\psi(s_1), \psi(s_2)) \\ \varphi_{\mathrm{dyn}}[C] &= \int_{s_1}^{s_2} \mathrm{d}s\, \mathrm{Im}(\psi(s), \dot{\psi}(s)). \end{aligned} \tag{11}$$

The term $\varphi_{\mathrm{p}}[C]$ – the subscript standing for Pancharatnam – is just the total phase seen in eq. (6). It depends only on the end points of $\mathcal{C}$ and is in that sense a nonlocal quantity. The term $\varphi_{\mathrm{dyn}}[C]$ is the 'dynamical' piece – it is an integral along $\mathcal{C}$ of a locally defined integrand.

Given any smooth curve C in ray space, we may choose *any* lift $\mathcal{C}$ in $\mathcal{H}$ projecting onto C. Then the calculation of the quantities $\varphi_{\mathrm{p}}[C]$ and $\varphi_{\mathrm{dyn}}[C]$ dependent on $\mathcal{C}$ is immediate. Their difference however is independent of the particular lift $\mathcal{C}$ and depends on C alone. In all this we see how easy it is to deal with open curves in Hilbert or ray space-noncyclic evolution. Whereas $\varphi_{\mathrm{dyn}}[C]$ is the integral along $\mathcal{C}$ of a suitably defined one-form on $\mathcal{H}$, $\varphi_g[C]$ cannot be displayed in any similar fashion at ray space level.

Particular kinds of lifts $C \to \mathcal{C}$ can lead to corresponding simplifications:

(i) Horizontal lifts: Here $(\psi(s), \dot{\psi}(s)) = 0$ throughout, so one has

$$\varphi_{\mathrm{dyn}}[C] = 0, \quad \varphi_g[C] = \varphi_{\mathrm{p}}[C]. \tag{12}$$

(ii) Pancharatnam lifts: Now the end points are chosen to be 'in phase' in Pancharatnam sense, so that $(\psi(s_1),\ \psi(s_2))$ is real positive, and then

$$\varphi_p[C] = 0,$$
$$\varphi_g[C] = -\varphi_{dyn}[C].$$

One can ask for an expression for $\varphi_g[C]$ directly in terms of density matrices, involving state vectors at all. One such expression which shows clearly the 'nonlc dependence of $\varphi_g[C]$ on C is

$$\varphi_g[C] = \arg \operatorname{Tr}\left\{\rho(s_1)\, P\left(\exp \int_{s_1}^{s_2} \mathrm{d}s\, \frac{\mathrm{d}\rho(s)}{\mathrm{d}s}\right)\right\}.$$

Reparametrization invariance and the geometric nature are evident.

All the earlier treatments and special cases can be easily recovered from formalism. In the case where one is dealing with some solution of the Schrödi equation for some given Hamiltonian, one can clearly pose and answer interes questions: how does the geometric phase behave under static and dynamic symmetrie the Hamiltonian and under antiunitary time reversal? It turns out to be inva under both static symmetries and time reversal, but in general not under dyna symmetries [8].

5. Connection to the Bargmann invariants

Given any continuous curve C of unit vectors in $\mathcal{H}$, with image C in ray space, possible to define the 'length' of C as a functional of C. By extremising this functi one arrives at the concept of geodesics in ray space. For convenience, *any* lift of a space geodesic will be regarded as a geodesic in Hilbert space as well. Then after suit choices of phases and parametrization, it turns out that the most general geodesic in describable as follows [8]

$$\psi(s) = \phi_1 \cos s + \phi_2 \sin s,$$
$$(\phi_1, \phi_1) = (\phi_2, \phi_2) = 1, \quad (\phi_1, \phi_2) = 0.$$

So it is basically a *real plane curve*, an arc of a circle in two dimensions. We see phases and the complex nature of $\mathcal{H}$ are all absent here. As a consequence we obtai key result [8]

$$\varphi_g[\text{geodesic } C \text{ in ray space}] = 0.$$

We now explain how this may be exploited.

If $\psi_1, \psi_2, \ldots, \psi_n$ are any n unit vectors in $\mathcal{H}$, given in this sequence, they define Bargmann invariant [9]

$$\Delta_n(\psi_1, \psi_2, \ldots, \psi_n) = (\psi_1, \psi_2)(\psi_2, \psi_3) \cdots (\psi_{n-1}, \psi_n)(\psi_n, \psi_1)$$
$$= \operatorname{Tr}(\rho_1 \rho_2 \cdots \rho_n),$$
$$\rho_1 = \psi_1 \psi_1^\dagger, \quad \rho_2 = \psi_2 \psi_2^\dagger, \ldots.$$

The context in which Bargmann introduced such expressions (for the case $n = 3$) was a discussion of the Wigner unitary–antiunitary theorem [10]: any symmetry in quantum mechanics, i.e., any probability preserving map of (pure state) rays on to rays, can be lifted to either a linear unitary or an antilinear antiunitary map at the Hilbert space level. The behaviour of Δ_3 distinguishes between these two alternatives, since it is invariant in the unitary case and goes to its complex conjugate in the antiunitary case.

The link between geometric phases and Bargmann invariants arises from the result (16) for geodesics. Given the points or vertices $\psi_1, \psi_2, \ldots, \psi_n$ needed to set up Δ_n, join ψ_1 to ψ_2 by a geodesic arc, ψ_2 to ψ_3 by another geodesic arc, and so on all the way till the geodesic arc from ψ_n back to ψ_1. This gives us an n-sided polygon in ray space as well, with $\rho_1, \rho_2, \ldots, \rho_n$ for vertices. We then find

$$\begin{aligned} &\varphi_g \quad [n\text{-sided polygon with vertices } \rho_1, \rho_2, \ldots, \rho_n] \\ &\quad = -\arg \Delta_n(\psi_1, \psi_2, \ldots, \psi_n). \end{aligned} \tag{18}$$

This is a nice connection between something old and something new, and it also gives some feeling for the nature of the geometric phase. It can and has been used in several ways: for instance the original definitions (10, 11) for $\varphi_g[C]$ can be recovered by a limiting process via a polygonal approximation to C, showing how the pieces $\varphi_p[C]$ and $\varphi_{dyn}[C]$ arise respectively from the last and the previous factors in $\Delta_n(\psi_1, \psi_2, \ldots, \psi_n)$; it has been shown that the very well-known Guoy phase in classical beam optics is an instance of the geometric phase, as seen via the Bargmann invariant [11].

6. The case of Lie group representations

The next major application of this formalism is to a study of geometric phases arising out of unitary Lie group representations [12]. Many special features are present on account of the algebraic and differential geometric properties that now become available.

The framework we use is the following. We have a (compact or noncompact) Lie group G, with a faithful unitary representation $\mathcal{U}(\cdot)$ of it acting on a Hilbert space $\mathcal{H}$. The representation need not be irreducible. Denote its hermitian generators by $\{T_r\}$. We have a given fiducial vector $\psi_0 \in \mathcal{H}$. We are interested in calculating geometric phases for smooth curves C which start out at ψ_0 and are produced by continuous group action on ψ_0. It turns out that the dynamical part of the geometric phase can be studied in great detail and brought to a maximally simplified form. We indicate the structures and results briefly.

Given ψ_0 and the unitary representation $\mathcal{U}(g)$ of G on $\mathcal{H}$, two subgroups of G are naturally defined

$$\begin{aligned} H_0 &= \{g \in G | \mathcal{U}(g)\, \psi_0 = \psi_0\} \\ &= \text{stability group of } \psi_0; \\ H &= \{g \in G | \mathcal{U}(g)\psi_0 = (\text{phase factor})\psi_0\} \\ &= \text{stability group of } \psi_0 \text{ up to phases.} \end{aligned} \tag{19}$$

Clearly H_0 is an invariant subgroup of H. For definiteness we make the assumption that H_0 is compact – this is so in all physically interesting cases.

Now the curves $\mathcal{C}$ that we are interested in can be viewed or described in two ways: either as lying in the orbit $\mathcal{O}(\psi_0)$ of ψ_0 in $\mathcal{H}$, which consists of the collection of vectors $\mathcal{U}(g)\,\psi_0$ for all $g \in G$; or as lying in the coset space $M_0 = G/H_0$. This is because by a well-known argument these two objects are the same – there is a one-to-one correspondence between vectors $\psi \in \mathcal{O}(\psi_0)$ and cosets in G/H_0. Hence

$$\begin{aligned}\mathcal{C} &= \{\mathcal{U}(g(s))\psi_0 | g(s) \text{a smooth curve in } G\} \\ &= \text{curve in } G - \text{orbit } \mathcal{O}(\psi_0) \text{ of } \psi_0 \\ &= \text{curve in coset space } M_0 = G/H_0. \end{aligned} \tag{20}$$

To calculate the dynamical phase $\varphi_{\text{dyn}}[\mathcal{C}]$, we need to appreciate that there are three (and only three) possible relations between H_0 and H, described as follows:

$$\begin{aligned} &\text{A)} \qquad H = H_0, \quad H/H_0 = \text{trivial}; \\ &\text{B)} \qquad H/H_0 = \text{discrete, nontrivial}; \\ &\text{C)} \qquad H/H_0 = U(1). \end{aligned} \tag{21}$$

In case (A) we are unable to change the phase of ψ_0 by group action, so $\mathcal{O}(\psi_0)$ consists of exactly one vector each from a certain collection of rays. In case (B) we can change the phase of ψ_0 by group action but only by certain discrete amounts, so $\mathcal{O}(\psi_0)$ consists of a discrete set of vectors drawn from each of a certain collection of rays. In case (C) we can change the phase of ψ_0 by any amount by suitable group action, so $\mathcal{O}(\psi_0)$ consists of a certain collection of entire rays. At the Lie algebra level, in cases (A) and (B), the subgroups H_0 and H have the same Lie algebra; while in case (C), H has an extra $U(1)$ generator invariant under H_0.

Now we turn to $\varphi_{\text{dyn}}[\mathcal{C}]$. By using the Wigner–Eckart theorem of quantum mechanics, and following also the spirit of that theorem, we find we can simplify $\varphi_{\text{dyn}}[\mathcal{C}]$ a great deal, and effect a neat separation of its algebraic and its geometric parts – namely, dependences on ψ_0, $\mathcal{U}(\cdot)$ on the one hand and dependence on $\mathcal{C}$ on the other. For this some standard differential geometric properties of G and M_0 are needed [13]. The Lie group G carries two sets of vector fields and their dual one forms, one set being left invariant and the other right invariant. These forms are the Maurer–Cartan one-forms and the Lie algebra structure can be expressed using either set. When we descend from G to the coset space $M_0 = G/H_0$, only some of the Maurer–Cartan forms survive and yield globally defined one-forms on M_0. These are the one-forms going with those generators of G which are scalar with respect to H_0. Using all this machinery and the standard Wigner–Eckart theorem we find

$$\varphi_{\text{dyn}}[\mathcal{C}] = (\psi_0, T_\rho \psi_0) \int_{\mathcal{C} \subset M_0} \widehat{\theta}^\rho \,. \tag{22}$$

There is a sum here on the repeated index ρ : $\{T_\rho\}$ is a complete independent set of H-scalar generators among $\{T_r\}$, and $\widehat{\theta}^\rho$ are the pull-backs to M_0 of the Maurer–Cartan forms on G that are associated with the T_ρ. Among the $\{T_\rho\}$ we may omit any generators of H_0 since they annihilate ψ_0; and the $\widehat{\theta}^\rho$ are globally well-defined on M_0.

We see that $\varphi_{\text{dyn}}[\mathcal{C}]$ is generally a sum of terms in each of which there is a neat separation of the algebraic representation dependent part from the geometric

dependence on C. This simplified expression allows a systematic study and comparison of cases where $U(g)$ may change but H_0 and C do not; when two groups G and G' are locally isomorphic but globally different, and the relevant coset spaces M_0, M_0' coincide, and so on.

Several applications of this formalism have been made: the Guoy phase was already mentioned, and this involves the metaplectic group $Mp(2)$; the case of $SU(3)$ representations including three-level quantum systems [14]; squeezing transformations etc. Indeed in practically every experimental realisation of geometric phases, it does happen that some Lie group is intimately involved!

7. Concluding remarks

This account conveys the value of viewing the geometric phase as a purely kinematic construct, and gives a new perspective on it. It has led to interesting questions and applications which are conceptually clear when expressed in the present framework. Lastly, as Feynman expressed himself in a well known paper: 'there is a pleasure in recognizing old things from a new point of view'.

References

[1] M V Berry, *Proc. R. Soc. (London)* **A392**, 45 (1984)

[2] S M Rytov, *Dokl. Akad. Nauk. (USSR)* **28**, 263 (1938)
V V Vladimirskii, *Dokl. Akad. Nauk (USSR)* **31**, 31 (1941)
S Pancharatnam, *Proc. Indian Acad. Sci.* **A44**, 247 (1956)

[3] See, for instance, A Shapere and F Wilczek, *Geometric phases in physics* (World Scientific Publishing Co. Pte. Ltd., Singapore, 1989)

[4] N Mukunda and R Simon, *Ann. Phys. (N.Y.)* **228**, 205, 269 (1993)

[5] H Weyl, *The theory of groups and quantum mechanics* (Dover Publications Inc., 1931) p. 4, 20

[6] See any standard text on quantum mechanics such as, for instance, L I Schiff, *Quantum Mechanics* (McGraw-Hill Book Company Inc., New York, 2nd edition, 1955) p. 213

[7] Y Aharonov and J Anandan, *Phys. Rev. Lett.* **58**, 1593 (1987)
J Samuel and R Bhandari, *Phys. Rev. Lett.* **60**, 2339 (1988)

[8] See the first of refs. [4] above

[9] V Bargmann, *J. Math. Phys.* **5**, 862 (1964)

[10] E P Wigner, *Group theory and its application to the quantum mechanics of atomic spectra* (Academic Press, New York, 1959) appendix to Ch. 20

[11] R Simon and N Mukunda, *Phys. Rev. Lett.* **70**, 880 (1993)

[12] See the second of refs. [4] above

[13] See, for instance, S Helgason, *Differential geometry and symmetric spaces* (Academic Press, New York, 1962)
B Schutz, *Geometrical methods of mathematical physics* (Cambridge University Press, London, 1980)
C J Isham, *Modern differential geometry for physicists* (World Scientific Publishing Co. Pte. Ltd., Singapore, 1989)

[14] G Khanna, S Mukhopadhyay, R Simon and N Mukunda, *Ann. Phys. (N.Y.)* **253**, 55 (1997)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 41–64

Supersymmetry in quantum mechanics

AVINASH KHARE
Institute of Physics, Bhubaneswar 751 005, India

Abstract. In the past ten years, the ideas of supersymmetry have been profitably applied to many nonrelativistic quantum mechanical problems. In particular, there is now a much deeper understanding of why certain potentials are analytically solvable. In this lecture I review the theoretical formulation of supersymmetric quantum mechanics and discuss many of its applications. I show that the well-known exactly solvable potentials can be understood in terms of a few basic ideas which include supersymmetric partner potentials and shape invariance. The connection between inverse scattering, isospectral potentials and supersymmetric quantum mechanics is discussed and multi-soliton solutions of the KdV equation are constructed. Further, it is pointed out that the connection between the solutions of the Dirac equation and the Schrödinger equation is exactly same as between the solutions of the MKdV and the KdV equations.

Keywords. Supersymmetry; shape invariant potentials; solvable potentials.

PACS No. 03.65

1. Introduction

Physicists have long strived to obtain a unified description of all basic interactions of nature, i.e. strong, electroweak, and gravitational interactions. Several ambitious attempts have been made in the last two decades, and it is now widely felt that supersymmetry (SUSY) is a necessary ingredient in any unifying approach. SUSY relates bosonic and fermionic degrees of freedom and has the virtue of taming ultraviolet divergences. One of the important predictions of SUSY theories is the existence of SUSY partners of quarks, leptons and gauge bosons. Despite the beauty of all these unified theories, there has so far been no experimental evidence of SUSY being realized in nature.

However, over the last 10 years, the ideas of SUSY have stimulated new approaches to other branches of physics [1] like nuclear, atomic, condensed matter, statistical physics as well as in quantum mechanics (QM). I have been fortunate to be involved in some of these developments in the area of supersymmetric quantum mechanics [2, 3]. Recently, Cooper, Sukhatme and myself have written an exhaustive Physics Reports on this topic where we have discussed many of these developments at length [4]. Today I would like to raise some of the issues in which SUSY has given us new insight in QM and discuss few of them in some detail.

1. It is well known that the infinite square well is one of the simplest exactly solvable problem in nonrelativistic QM and the energy eigenvalues are given by $E_n = c(n+1)^2$ where c is constant. Are there other potentials for which the energy eigenvalues have a similar form and is there a simple way of obtaining these potentials?

2. Free particle is obviously the simplest example in QM with no bound states, no reflection and transmission probability being unity. Are there other (nontrivial) potentials for which also there is no reflection and is it possible to easily construct such potentials?
3. Why is Schrödinger equation analytically solvable in the case of few potentials? Another question is if the one dimensional harmonic oscillator the only potential which can be solved by operator method? In this context, recall that the operator method of solving the one dimensional problem is in fact the whole basis of quantum field theory as well as many body theory.
4. It is well known that given a potential $V(x)$, the corresponding energy eigenvalues E_n, and the scattering matrix (the reflection and transmission coefficients $R(k)$ and $T(k)$ in the one dimensional case or phase shifts in the three dimensional case) are unique. Is the converse also true i.e. given E_n, $R(k)$ and $T(k)$ is the corresponding potential unique? If not then how does one construct the various potentials with the same E_n, R and T?
5. A related question is about the construction of the soliton solutions of the KdV and other nonlinear equations. Can these be easily constructed from the formalism of SUSY QM?
6. What is the connection between the Dirac and the Schrödinger equations? In particular, knowing the solution of the Schrödinger problem does there always exist a corresponding exactly solvable Dirac problem and what is the precise connection between the two?
7. Is there a unified treatment for constructing the bound states in the (classical) continuum?
8. Are there semiclassical approximations which do even better than the usual WKB approximation? For example is there an approximation scheme for which the lowest order is exact while all higher order corrections are zero?
9. Finally, can one also analytically solve few noncentral potentials by using operator method alone?

Before I discuss in some detail as to what SUSY QM has to say about these questions, I shall briefly discuss the formalism of SUSY QM and show that because of the underline SUSY, the energy eigenvalues, the eigenfunctions and the S-matrix of the two partner potentials are related in a very definite way.

2. Formalism

One of the key ingredients in solving exactly for the spectrum of one dimensional potential problems is the connection between the bound state wave functions and the potential. It is not usually appreciated that once one knows the ground state wave function (or any other bound state wave function) then one knows exactly the potential (up to a constant). Let us choose the ground state energy for the moment to be zero. Then one has from the Schrödinger equation that the ground state wave function $\psi_0(x)$ obeys [4]

$$H_1\psi_0(x) = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi_0}{\mathrm{d}x^2} + V_1(x)\psi_0(x) = 0, \tag{1}$$

so that

$$V_1(x) = \frac{\hbar^2}{2m}\frac{\psi_0''(x)}{\psi_0(x)}. \tag{2}$$

This allows a global reconstruction of the potential $V_1(x)$ from a knowledge of its ground state wave function which has no nodes. Once we realize this, it is now very simple to factorize the Hamiltonian using the following ansatz:

$$H_1 = A^\dagger A, \tag{3}$$

where

$$A = \frac{\hbar}{\sqrt{2m}}\frac{\mathrm{d}}{\mathrm{d}x} + W(x), \quad A^\dagger = \frac{-\hbar}{\sqrt{2m}}\frac{\mathrm{d}}{\mathrm{d}x} + W(x). \tag{4}$$

This allows us to identify

$$V_1(x) = W^2(x) - \frac{\hbar}{\sqrt{2m}}W'(x). \tag{5}$$

This equation is the well-known Riccati equation. The quantity $W(x)$ is generally referred to as the 'superpotential' in SUSY QM literature. The solution for $W(x)$ in terms of the ground state wave function is

$$W(x) = -\frac{\hbar}{\sqrt{2m}}\frac{\psi_0'(x)}{\psi_0(x)}. \tag{6}$$

This solution is obtained by recognizing that once we satisfy $A\psi_0 = 0$, we automatically have a solution to $H_1\psi_0 = A^\dagger A\psi_0 = 0$.

The next step in constructing the SUSY theory related to the original Hamiltonian H_1 is to define the operator $H_2 = AA^\dagger$ obtained by reversing the order of A and $A^\dagger$. A little simplification shows that the operator H_2 is in fact a Hamiltonian corresponding to a new potential $V_2(x)$

$$H_2 = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}X^2} + V_2(x), \quad V_2(x) = W^2(x) + \frac{\hbar}{\sqrt{2m}}W'(x). \tag{7}$$

The potentials $V_1(x)$ and $V_2(x)$ are known as supersymmetric partner potentials.

As we shall see, the energy eigenvalues, the wave functions and the S-matrices of H_1 and H_2 are related. To that end notice that the energy eigenvalues of both H_1 and H_2 are positive semi-definite ($E_n^{(1,2)} \geq 0$) . For $n > 0$, the Schrödinger equation for H_1

$$H_1\psi_n^{(1)} = A^\dagger A\psi_n^{(1)} = E_n^{(1)}\psi_n^{(1)} \tag{8}$$

implies

$$H_2(A\psi_n^{(1)}) = AA^\dagger A\psi_n^{(1)} = E_n^{(1)}(A\psi_n^{(1)}). \tag{9}$$

Similarly, the Schrödinger equation for H_2

$$H_2\psi_n^{(2)} = AA^\dagger\psi_n^{(2)} = E_n^{(2)}\psi_n^{(2)} \tag{10}$$

implies

$$H_1(A^\dagger\psi_n^{(2)}) = A^\dagger AA^\dagger\psi_n^{(2)} = E_n^{(2)}(A^\dagger\psi_n^{(2)}). \tag{11}$$

From eqs (8)–(11) and the fact that $E_0^{(1)} = 0$, it is clear that the eigenvalues and eigenfunctions of the two Hamiltonians H_1 and H_2 are related by $(n = 0, 1, 2, \ldots)$

$$E_n^{(2)} = E_{n+1}^{(1)}, \quad E_0^{(1)} = 0, \tag{12}$$

$$\psi_n^{(2)} = [E_{n+1}^{(1)}]^{-1/2} A\psi_{n+1}^{(1)}, \tag{13}$$

$$\psi_{n+1}^{(1)} = [E_n^{(2)}]^{-1/2} A^\dagger\psi_n^{(2)}. \tag{14}$$

Notice that if $\psi_{n+1}^{(1)}$ $(\psi_n^{(2)})$ of H_1 (H_2) is normalized then the wave function $\psi_n^{(2)}$ $(\psi_{n+1}^{(1)})$ in eqs (13) and (14) is also normalized. Further, the operator A $(A^\dagger)$ not only converts an eigenfunction of H_1 (H_2) into an eigenfunction of $H_2(H_1)$ with the same energy, but it also destroys (creates) an extra node in the eigenfunction. Since the ground state wave function of H_1 is annihilated by the operator A, this state has no SUSY partner. Thus the picture we get is that knowing all the eigenfunctions of H_1 we can determine the eigenfunctions of H_2 using the operator A, and vice versa using $A^\dagger$ we can reconstruct all the eigenfunctions of H_1 from those of H_2 except for the ground state.

The underlying reason for the degeneracy of the spectra of H_1 and H_2 can be understood most easily from the properties of the SUSY algebra. That is we can consider a matrix SUSY Hamiltonian of the form

$$H = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \tag{15}$$

which contains both H_1 and H_2. This matrix Hamiltonian is part of a closed algebra which contains both bosonic and fermionic operators with commutation and anti-commutation relations. We consider the operators

$$Q = \begin{bmatrix} 0 & 0 \\ A & 0 \end{bmatrix}, \tag{16}$$

$$Q^\dagger = \begin{bmatrix} 0 & A^\dagger \\ 0 & 0 \end{bmatrix} \tag{17}$$

in conjunction with H. The following commutation and anticommutation relations then describe the closed superalgebra $sl(1/1)$:

$$[H, Q] = [H, Q^\dagger] = 0,$$
$$\{Q, Q^\dagger\} = H, \quad \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0. \tag{18}$$

The fact that the supercharges Q and $Q^\dagger$ commute with H is responsible for the degeneracy.

Supersymmetry also allows one to relate the reflection and transmission coefficients in situations where the two partner potentials have continuum spectra. In order for scattering

to take place in both of the partner potentials, it is necessary that the potentials $V_{1,2}$ are finite as $x \to -\infty$ or as $x \to +\infty$ or both. Define

$$W(x \to \pm\infty) = W_{\pm}. \tag{19}$$

Then

$$V_{1,2} \to W_{\pm}^2 \quad \text{as} \quad x \to \pm\infty. \tag{20}$$

Let us consider an incident plane wave e^{ikx} of energy E coming from the direction $x \to -\infty$. As a result of scattering from the potentials $V_{1,2}(x)$ one would obtain transmitted waves $T_{1,2}(k)e^{ik'x}$ and reflected waves $R_{1,2}(k)e^{-ikx}$. Thus we have

$$\begin{aligned}\psi^{(1,2)}(k, x \to -\infty) &\to e^{ikx} + R_{1,2}e^{-ikx},\\ \psi^{(1,2)}(k', x \to +\infty) &\to T_{1,2}e^{ik'x}.\end{aligned} \tag{21}$$

SUSY connects continuum wave functions of H_1 and H_2 having the same energy analogously to what happens in the discrete spectrum. Thus we have the relationships:

$$\begin{aligned}e^{ikx} + R_1e^{-ikx} &= N[(-ik + W_-)e^{ikx} + (ik + W_-)e^{-ikx}R_2],\\ T_1e^{ik'x} &= N[(-ik' + W_+)e^{ik'x}T_2],\end{aligned} \tag{22}$$

where N is an overall normalization constant. On equating terms with the same exponent and eliminating N, we find

$$\begin{aligned}R_1(k) &= \left(\frac{W_- + ik}{W_- - ik}\right)R_2(k),\\ T_1(k) &= \left(\frac{W_+ - ik'}{W_- - ik}\right)T_2(k),\end{aligned} \tag{23}$$

where k and k' are given by

$$k = (E - W_-^2)^{1/2}, \quad k' = (E - W_+^2)^{1/2}. \tag{24}$$

A few remarks are now in order at this stage.

(1) Clearly $|R_1|^2 = |R_2|^2$ and $|T_1|^2 = |T_2|^2$, that is the partner potentials have identical reflection and transmission probabilities.
(2) $R_1(T_1)$ and $R_2(T_2)$ have the same poles in the complex plane except that $R_1(T_1)$ has an extra pole at $k = -iW_-$. This pole is on the positive imaginary axis only if $W_- < 0$ in which case it corresponds to a zero energy bound state.
(3) In the special case that $W_+ = W_-$, we have that $T_1(k) = T_2(k)$.
(4) When $W_- = 0$ then $R_1(k) = -R_2(k)$.

3. SUSY partner of the infinite square well

Let us look at a well known potential, namely the infinite square well and determine its SUSY partner potential. Consider a particle of mass m in an infinite square well potential

of width L.

$$V(x) = 0, \quad 0 \leq x \leq L,$$
$$= \infty, \quad -\infty < x < 0, x > L. \tag{25}$$

The ground state wave function is known to be

$$\psi_0^{(1)} = (2/L)^{1/2} \sin(\pi x/L), \quad 0 \leq x \leq L, \tag{26}$$

and the ground state energy is $E_0 = \hbar^2\pi^2/2mL^2$.

Subtracting off the ground state energy so that we can factorize the Hamiltonian we have for $H_1 = H - E_0$ that the energy eigenvalues are

$$E_n^{(1)} = \frac{n(n+2)}{2mL^2}\hbar^2\pi^2 \tag{27}$$

and the eigenfunctions are

$$\psi_n^{(1)} = (2/L)^{1/2} \sin\frac{(n+1)\pi x}{L}, \quad 0 \leq x \leq L. \tag{28}$$

The superpotential for this problem is readily obtained using eq. (6)

$$W(x) = -\frac{\hbar}{\sqrt{2m}}\frac{\pi}{L}\cot(\pi x/L) \tag{29}$$

and hence the supersymmetric partner potential V_2 is

$$V_2(x) = \frac{\hbar^2\pi^2}{2mL^2}[2\mathrm{cosec}^2(\pi x/L) - 1]. \tag{30}$$

The wave functions for H_2 are obtained by applying the operator A to the wave functions of H_1. In particular we find that

$$\psi_0^{(2)} \propto \sin^2(\pi x/L), \quad \psi_1^{(2)} \propto \sin(\pi x/L)\sin(2\pi x/L). \tag{31}$$

Thus we have shown using SUSY that two rather different potentials corresponding to H_1 and H_2 have exactly the same spectra except for the fact that H_2 has one fewer bound state.

4. Reflectionless potentials

It is clear from the discussion in § 2 that if one of the partner potentials is a constant potential (i.e. a free particle), then the other partner will be of necessity reflectionless. In this way we can understand the reflectionless potentials of the form $V(x) = A\,\mathrm{sech}^2\alpha x$ which play a critical role in understanding the soliton solutions of the KdV hierarchy. Let us consider the superpotential

$$W(x) = A\tanh\alpha x. \tag{32}$$

The two partner potentials are

$$V_1 = A^2 - A\left(A + \alpha\frac{\hbar}{\sqrt{2m}}\right)\operatorname{sech}^2\alpha x,$$

$$V_2 = A^2 - A\left(A - \alpha\frac{\hbar}{\sqrt{2m}}\right)\operatorname{sech}^2\alpha x. \tag{33}$$

We see that for $A = \alpha(\hbar/\sqrt{2m})$, $V_2(x)$ corresponds to a constant potential so that the corresponding V_1 is a reflectionless potential. It is worth noting that V_1 is $\hbar$-dependent. One can in fact rigorously show, though it is not mentioned in most text books, that the reflectionless potentials are necessarily $\hbar$-dependent. Proceeding in the same way one can obtain reflectionless potentials with n bound states in case $A = n\alpha(\hbar/\sqrt{2m})$, $(n = 2, 3, \ldots;$ note that V_1 has precisely one bound state).

Recently we [5, 6] have also shown how to construct reflectionless potentials with infinite number of bound states.

It is worth mentioning here that these reflectionless potentials are intimately related with the soliton solutions of the KdV equation

$$u_t = 6uu_x - u_{xxx}. \tag{34}$$

We shall have something more to say about this question in a later section.

5. Shape invariance and solvable potentials

Most text books on quantum mechanics describe how the one dimensional harmonic oscillator problem can be elegantly solved using the raising and lowering operator method. Using the ideas of SUSY QM developed in § 2 and an integrability condition called the shape invariance condition [8], we now show that the operator method for the harmonic oscillator can be generalized to the whole class of shape invariant potentials (SIP) which include all the popular, analytically solvable potentials. Indeed, we shall see that for such potentials, the generalized operator method quickly yields all the bound state energy eigenvalues, eigenfunctions as well as the scattering matrix. It turns out that this approach is essentially equivalent to Schrödinger's method of factorization [7] although the language of SUSY is more appealing.

Let us now explain precisely what one means by shape invariance. If the pair of SUSY partner potentials $V_{1,2}(x)$ defined in § 2 are similar in shape and differ only in the parameters that appear in them, then they are said to be shape invariant. More precisely, if the partner potentials $V_{1,2}(x; a_1)$ satisfy the condition

$$V_2(x; a_1) = V_1(x; a_2) + R(a_1), \tag{35}$$

where a_1 is a set of parameters, a_2 is a function of a_1 (say $a_2 = f(a_1)$) and the remainder $R(a_1)$ is independent of x, then $V_1(x; a_1)$ and $V_2(x; a_1)$ are said to be shape invariant. The shape invariance condition (35) is an integrability condition.

Let us start from the SUSY partner Hamiltonians H_1 and H_2 whose eigenvalues and eigenfunctions are related by SUSY. Further, since SUSY is unbroken we know that

$$E_0^{(1)}(a_1) = 0, \quad \psi_0^{(1)}(x; a_1) = N\exp\left[-\int^x W_1(y; a_1)\mathrm{d}y\right]. \tag{36}$$

We now show that the entire spectrum of H_1 can be very easily obtained algebraically by using the shape invariance condition (35). To that purpose, let us construct a series of Hamiltonians H_s, $s = 1, 2, 3 \ldots$. In particular, following the discussion of the last section it is clear that if H_1 has p bound states then one can construct p such Hamiltonians $H_1, H_2, \ldots, H_p$ and the nth Hamiltonian H_n will have the same spectrum as H_1 except that the first $n-1$ levels of H_1 will be absent in H_n. On repeatedly using the shape invariance condition (35), it is clear that

$$H_s = -\frac{d^2}{dx^2} + V_1(x; a_s) + \sum_{k=1}^{s-1} R(a_k), \tag{37}$$

where $a_s = f^{s-1}(a_1)$ i.e. the function f applied $s-1$ times. Let us compare the spectrum of H_s and H_{s+1}. In view of eqs (35) and (37) we have

$$\begin{aligned} H_{s+1} &= -\frac{d^2}{dx^2} + V_1(x; a_{s+1}) + \sum_{k=1}^{s} R(a_k) \\ &= -\frac{d^2}{dx^2} + V_2(x; a_s) + \sum_{k=1}^{s-1} R(a_k). \end{aligned} \tag{38}$$

Thus H_s and H_{s+1} are SUSY partner Hamiltonians and hence have identical bound state spectra except for the ground state of H_s whose energy is

$$E_0^{(s)} = \sum_{k=1}^{s-1} R(a_k). \tag{39}$$

This follows from eq. (37) and the fact that $E_0^{(1)} = 0$. On going back from H_s to H_{s-1} etc, we would eventually reach H_2 and H_1 whose ground state energy is zero and whose nth level is coincident with the ground state of the Hamiltonian H_n. Hence the complete eigenvalue spectrum of H_1 is given by

$$E_n^-(a_1) = \sum_{k=1}^{n} R(a_k); \quad E_0^-(a_1) = 0. \tag{40}$$

We now show that, similar to the case of the one dimensional harmonic oscillator, the bound state wave functions $\psi_n^{(1)}(x; a_1)$ for any shape invariant potential can also be easily obtained from its ground state wave function $\psi_0^{(1)}(x; a_1)$ which in turn is known in terms of the superpotential. This is possible because the operators A and $A^\dagger$ link up the eigenfunctions of the same energy for the SUSY partner Hamiltonians $H_{1,2}$. Let us start from the Hamiltonian H_s as given by eq. (37) whose ground state eigenfunction is then given by $\psi_0^{(1)}(x; a_s)$. On going from H_s to H_{s-1} to H_2 to H_1 and using eq. (14) we then find that the nth state unnormalized, energy eigenfunction $\psi_n^{(1)}(x; a_1)$ for the original Hamiltonian $H_1(x; a_1)$ is given by

$$\psi_n^{(1)}(x; a_1) \propto A^\dagger(x; a_1) A^\dagger(x; a_2) \cdots A^\dagger(x; a_n) \psi_0^{(1)}(x; a_{n+1}), \tag{41}$$

which is clearly a generalization of the operator method of constructing the energy eigenfunctions for the one dimensional harmonic oscillator.

Finally, it is worth noting that in view of the shape invariance condition (35), the relation (23) between scattering amplitudes takes a particularly simple form

$$R_1(k;a_1) = \left(\frac{W_-(a_1)+ik}{W_-(a_1)-ik}\right)R_1(k;a_2), \tag{42}$$

$$T_1(k;a_1) = \left(\frac{W_+(a_1)-ik'}{W_-(a_1)-ik}\right)T_1(k;a_2), \tag{43}$$

thereby relating the reflection and transmission coefficients of the same Hamiltonian H_1 at a_1 and $a_2(=f(a_1))$.

Let us now discuss the interesting question of the classification of various solutions to the shape invariance condition (35). This is clearly an important problem because once such a classification is available, then one can discover new SIPs which are solvable by purely algebraic methods. Although the general problem is still unsolved, two classes of solutions have been found and discussed. In the first class, the parameters a_1 and a_2 are related to each other by translation $(a_2 = a_1 + \alpha)$ [9]. Remarkably enough, all well known analytically solvable potentials found in most textbooks on nonrelativistic quantum mechanics belong to this class. In table 1, we give expressions for the various shape invariant potentials $V_1(x)$, superpotentials $W(x)$, parameters a_1 and a_2 and the corresponding energy eigenvalues $E_n^{(1)}$ and eigenfunctions $\psi_n^{(1)}$ [10, 11].

As an illustration, let us consider the superpotential $(\hbar = 2m = 1)$

$$W(x;A,B) = A\tan\alpha x - B\cot\alpha x; \quad A,B>0. \tag{44}$$

The corresponding partner potentials are

$$\begin{aligned} V_1(x;A,B) &= -(A+B)^2 + A(A-\alpha)\sec^2\alpha x + B(B-\alpha)\mathrm{cosec}^2\alpha x, \\ V_2(x;A,B) &= -(A+B)^2 + A(A+\alpha)\sec^2\alpha x + B(B+\alpha)\mathrm{cosec}^2\alpha x. \end{aligned} \tag{45}$$

V_1 and V_2 are often called Pöschl–Teller I potentials in the literature. They are shape invariant partner potentials since

$$V_2(x;A,B) = V_1(x;A+\alpha,B+\alpha) + (A+B+2\alpha)^2 - (A+B)^2 \tag{46}$$

and in this case

$$\{a_1\} = (A,B); \{a_2\} = (A+\alpha, B+\alpha), R(a_1) = (A+B+2\alpha)^2 - (A+B)^2. \tag{47}$$

In view of eq. (40), the bound state energy eigenvalues of the potential $V_1(x;A,B)$ are then given by

$$E_n^{(1)} = \sum_{k=1}^{n} R(a_k) = (A+B+2n\alpha)^2 - (A+B)^2. \tag{48}$$

The ground state wave function of $V_1(x;A,B)$ is calculated from the superpotential W as given by eq. (44). We find

$$\psi_0^{(1)}(x;A,B) \propto (\cos\alpha x)^s(\sin\alpha x)^\lambda, \tag{49}$$

Table 1. All known shape invariant potentials in which the parameters a_2 and a_1 are related by translation ($a_2 = a_1 + \alpha$). Unless stated otherwise, all constants like A, B, α, ω etc are all taken $\geqslant 0$ while $\hbar$ and $2m$ are taken to be 1.

Name of potential	Super potential $W(x)$	Potential $V_1(x; a_1)$	a_1	a_2	Eigenvalue $E_n^{(1)}$	Variable y	Wave function $\psi_n(y)$
Shifted oscillator	$\frac{1}{2}\omega x - b$	$\frac{1}{4}\omega^2\left(x - \frac{2b}{\omega}\right)^2 - \omega/2$	ω	ω	$n\omega$	$y = \left(\frac{1}{2}\omega\right)^{1/2} \times \left(x - \frac{2b}{\omega}\right)$	$\exp\left(-\frac{1}{2}y^2\right) \times H_n(y)$
Three dimensional oscillator	$\frac{1}{2}\omega r - \frac{(l+1)}{r}$	$\frac{1}{4}\omega^2 r^2 + \frac{l(l+1)}{r^2}$ $-(l+3/2)\omega$	l	$l+1$	$2n\omega$	$y = \frac{1}{2}\omega r^2$	$y^{(l+1)/2}\exp\left(-\frac{1}{2}y\right)$ $\times L_n^{l+1/2}(y)$
Coulomb	$\frac{e^2}{2(l+1)} - \frac{(l+1)}{r}$	$-\frac{e^2}{r} + \frac{l(l+1)}{r^2}$ $\frac{e^4}{4(l+1)^2}$	l	$l+1$	$\frac{e^2}{4}\left(\frac{1}{(l+1)^2}\right.$ $\left. - \frac{1}{(n+l+1)^2}\right)$	$y = \frac{re^2}{(n+l+1)}$	$y^{l+1}\exp\left(-\frac{1}{2}y\right)$ $\times L_n^{2l+1}(y)$
Morse	A $-B\exp(-\alpha x)$	$A^2 + B^2\exp(-2\alpha x)$ $-2B(A+\alpha/2)$ $\times \exp(-\alpha x)$	A	$A-\alpha$	$A^2 - (A - n\alpha)^2$	$y = (2B/\alpha)$ $\times \exp(-\alpha x), s = A/\alpha$	$y^{s-n}\exp\left(-\frac{1}{2}y\right)$ $\times L_n^{2s-2n}(y)$
Scarf II (hyperbolic)	$A\tanh\alpha x$ $+B\,\mathrm{sech}\,\alpha x$	$A^2 + (B^2 - A^2 - A\alpha)\mathrm{sech}^2\alpha x$ $+B(2A+\alpha)\mathrm{sech}\,\alpha x\tanh\alpha x$	A	$A-\alpha$	$A^2 - (A - n\alpha)^2$	$y = \sinh\alpha x,$ $s = A/\alpha, \lambda = B/\alpha$	$i^n(1+y^2)^{-s/2}$ $\times \exp(-\lambda\tan^{-1}y)$ $\times P_n^{(i\lambda - s - \frac{1}{2}, -i\lambda - s - \frac{1}{2})}(iy)$

Rosen–Morse II (hyperbolic)	$A\tanh\alpha x$ $+B/A$ $(B<A^2)$	A^2+B^2/A^2 $-A(A+\alpha)\text{sech}^2\alpha x$ $+2B\tanh\alpha x$	A	$A-\alpha$	$A^2-(A-n\alpha)^2$ $+\frac{B^2}{A^2}-\frac{B^2}{(A-n\alpha)^2}$	$y=\tanh\alpha x,$ $s=A/\alpha,\lambda=B/\alpha^2$ $a=\lambda/(s-n)$	$(1-y)^{(s-n+a)/2}$ $\times(1+y)^{(s-n-a)/2}$ $\times P_n^{(s-n+a,s-n-a)}(y)$
Eckart	$-A\coth\alpha r$ $+B/A$ $(B>A^2)$	A^2+B^2/A^2 $+A(A-\alpha)\text{cosech}^2\alpha r$ $-2B\coth\alpha r$	A	$A+\alpha$	$A^2-(A+n\alpha)^2$ $+\frac{B^2}{A^2}-\frac{B^2}{(A+n\alpha)^2}$	$y=\coth\alpha r,$ $s=A/\alpha,\lambda=B/\alpha^2,$ $a=\lambda/(n+s)$	$(y-1)^{-(s+n-a)/2}$ $\times(y+1)^{-(s+n+a)/2}$ $\times P_n^{(-s-n+a,-s-n-a)}(y)$
Scarf I (trigonometric)	$-A\tan\alpha x$ $+B\sec\alpha x$ $\left(-\frac{1}{2}\pi\le\alpha x\le\frac{1}{2}\pi\right)$	$-A^2+(A^2+B^2-A\alpha)\sec^2\alpha x$ $-B(2A-\alpha)\tan\alpha x\sec\alpha x$	A	$A+\alpha$	$(A+n\alpha)^2-A^2$	$y=\sin\alpha x,$ $s=A/\alpha,\lambda=B/\alpha$	$(1-y)^{(s-\lambda)/2}$ $\times(1+y)^{(s+\lambda)/2}$ $\times P_n^{(s-\lambda-1/2,s+\lambda-1/2)}(y)$
Generalized Pöschl–Teller	$A\coth\alpha r$ $-B\,\text{cosech}\,\alpha r$ $(A<B)$	$A^2+(B^2+A^2+A\alpha)\text{cosech}^2\alpha r$ $-B(2A+\alpha)\coth\alpha r\,\text{cosech}\,\alpha r$	A	$A-\alpha$	$A^2-(A-n\alpha)^2$	$y=\cosh\alpha r,$ $s=A/\alpha,\lambda=B/\alpha$	$(y-1)^{(\lambda-s)/2}$ $\times(y+1)^{-(\lambda+s)/2}$ $\times P_n^{(\lambda-s-\frac{1}{2},-\lambda-s-\frac{1}{2})}(y)$
Rosen–Morse I (trigonometric)	$-A\cot\alpha x$ $-B/A$ $0\le\alpha x\le\pi$	$A(A-\alpha)\,\text{cosec}^2\alpha x$ $+2B\cot\alpha x$ $-A^2+B^2/A^2$	A	$A+\alpha$	$(A+n\alpha)^2-A^2$ $+\frac{B^2}{A^2}-\frac{B^2}{(A+n\alpha)^2}$	$y=i\cot\alpha x,$ $s=A/\alpha,\lambda=B/\alpha^2,$ $a=\lambda/(s+n)$	$(y^2-1)^{-(s+n/2)}$ $\times\exp(a\alpha x)$ $\times P_n^{(-s-n+ia,-s-n-ia)}(y)$

where

$$s = A/\alpha; \quad \lambda = B/\alpha. \tag{50}$$

The requirement of $A, B > 0$ that we have assumed in eq. (44) guarantees that $\psi_0^{(1)}(x; A, B)$ is well behaved and hence acceptable as $x \longrightarrow 0, \pi/2\alpha$. Using this expression for the ground state wave function one can also obtain explicit expressions for the bound state eigenfunctions $\psi_n^{(1)}(x; A, B)$ [12].

Thus what we have shown here is that shape invariance is a sufficient condition for exact solvability. But is it also a necessary condition? This question has been discussed in detail in ref. [9] where it has been shown that the answer to the question is no.

For the SIPs (with translation) given in table 1, the reflection and transmission amplitudes $R_1(k)$ and $T_1(k)$ (or phase shift $\delta_1(k)$ for the three-dimensional case) can also be calculated by operator methods. Let us first notice that since for all the cases $a_2 = a_1 + \alpha$, hence $R_1(k; a_1)$ and $T_1(k; a_1)$ are determined for all values of a_1 from eqs (42) and (43) provided they are known in a finite strip. For example, let us consider the shape invariant superpotential

$$W = n \tanh x, \tag{51}$$

where n is positive integer $(1, 2, 3, \ldots)$. The two partner potentials

$$\begin{aligned} V_1(x; n) &= n^2 - n(n+1)\operatorname{sech}^2 x, \\ V_2(x; n) &= n^2 - n(n-1)\operatorname{sech}^2 x, \end{aligned} \tag{52}$$

are clearly shape invariant with

$$a_1 = n, \quad a_2 = n - 1. \tag{53}$$

On going from V_1 to V_2 to V_3 etc., we will finally reach the free particle potential which is reflectionless and for which $T = 1$. Thus we immediately conclude that the series of potentials $V_1, V_2, \ldots$ are all reflectionless and the transmission coefficient of the reflectionless potential $V_1(x; n)$ is given by

$$\begin{aligned} T_1(k, n) &= \frac{(n - ik)(n - 1 - ik) \cdots (1 - ik)}{(-n - ik)(-n + 1 - ik) \cdots (-1 - ik)} \\ &= \frac{\Gamma(-n - ik)\Gamma(n + 1 - ik)}{\Gamma(-ik)\Gamma(1 - ik).} \end{aligned} \tag{54}$$

The scattering amplitudes of the Coulomb and the potential corresponding to $W = A \tanh x + B \operatorname{sech} x$ have also been obtained in this way.

There is, however, a straightforward method [13] for calculating the scattering amplitudes by making use of the nth state wave functions as given in table 1. In order to impose boundary conditions appropriate to the scattering problem, two modifications of the bound state wave functions have to be made: (i) instead of the parameter n labelling the number of nodes, one must use the wave number k' so that the asymptotic behaviour is $\exp(ik'x)$ as $x \to \infty$, (ii) the second solution of the Schrödinger equation must be kept (it had been discarded for bound state problems since it diverged asymptotically).

In this way the scattering amplitude of all the SIPs of table 1 have been calculated in ref. [13].

For almost nine years, it was believed that the only shape invariant potentials are those given in table 1 and that there were no more shape invariant potentials. However, very recently we have been able to discover a huge class of new shape invariant potentials [5, 6]. It turns out that for many of these new shape invariant potentials, the parameters a_2 and a_1 are related by scaling ($a_2 = qa_1, 0 < q < 1$) rather than by translation, a choice motivated by the recent interest in q-deformed Lie algebras. We shall see that many of these potentials are reflectionless and have an infinite number of bound states. So far none of these potentials have been obtained in a closed form but are obtained only in a series form.

6. Shape invariance and noncentral solvable potentials

We have seen that using the ideas of SUSY and shape invariance, a number of potential problems can be solved algebraically. Most of these potentials are either one dimensional or are central potentials which are again essentially one dimensional but on the half line. It may be worthwhile to enquire if one can also algebraically solve some noncentral but separable potential problems. As has been shown recently [14], the answer to the question is yes. It turns out that the problem is algebraically solvable so long as the separated problems for each of the coordinates belong to the class of SIP. As an illustration, let us discuss noncentral separable potentials in spherical polar coordinates.

In spherical polar coordinates the most general potential for which the Schrödinger equation is separable is given by

$$V(r,\theta,\phi) = V(r) + \frac{V(\theta)}{r^2} + \frac{V(\phi)}{r^2\sin^2\theta}, \tag{55}$$

where $V(r), V(\theta)$ and $V(\phi)$ are arbitrary functions of their argument. First, let us see why the Schrödinger equation with a potential of the form given by eq. (55) is separable in the (r,θ,ϕ) coordinates. The equation for the wave function $\psi(r,\theta,\phi)$ is

$$\left[-\left(\frac{\partial^2\psi}{\partial^2 r} + \frac{2}{r}\frac{\partial\psi}{\partial r}\right) - \frac{1}{r^2}\left(\frac{\partial^2\psi}{\partial\theta^2} + \cot\theta\frac{\partial\psi}{\partial\theta}\right) - \frac{1}{r^2\sin^2\theta}\frac{\mathrm{d}^2\psi}{\partial\phi^2}\right] = (E-V)\psi. \tag{56}$$

It is convenient to write $\psi(r,\theta,\phi)$ as

$$\psi(r,\theta,\phi) = \frac{R(r)}{r}\frac{H(\theta)}{(\sin\theta)^{1/2}}K(\phi). \tag{57}$$

Substituting eq. (57) in eq. (56) and using the standard separation of variables procedure, one obtains the following equations for the functions $K(\phi), H(\theta)$ and $R(r)$:

$$-\frac{\mathrm{d}^2K}{\mathrm{d}\phi^2} + V(\phi)K(\phi) = m^2K(\phi), \tag{58}$$

$$-\frac{\mathrm{d}^2H}{\mathrm{d}\theta^2} + [V(\theta) + (m^2 - \tfrac{1}{4})\,\mathrm{cosec}^2\theta]H(\theta) = l^2H(\theta), \tag{59}$$

$$-\frac{d^2R}{dr^2}+\left[V(r)+\frac{(l^2-(1/4))}{r^2}\right]R(r)=ER(r), \tag{60}$$

where m^2 and l^2 are separation constants.

As has been shown in ref. [14], the three Schrödinger equations given by (58), (59) and (60) may be solved algebraically by choosing appropriate SIPs for $V(\phi), V(\theta)$ and $V(r)$. Details can be found in ref. [14].

Generalization of this technique to noncentral but separable potentials in other orthogonal curvilinear coordinate systems as well as in other dimensions is quite straightforward. Further, as we show now, one could use this trick to discover a number of new exactly solvable three-body potentials in one dimension.

Many years ago, in a classic paper, Calogero [15] gave the complete solution of the Schrödinger equation for three particles in one dimension, interacting via two-body harmonic and inverse-square potentials. Attention then shifted to the exact solutions of the many-body problem and the general question of integrability. Recently, there has been renewed interest in the one-dimensional many-body problem in connection with the physics of spin chains. Also, there has been a recent generalization of Calogero's potential for N-particles to SUSY QM. In fact one now believes that the Calogero–Sutherland model represents an example of ideal gas in one dimension with fractional statistics [16].

The important point to note is that three particles in one dimension, after the center of mass motion is eliminated, have two degrees of freedom. This may therefore be mapped on to a one-body problem in two dimensions. Calogero [15] considered the case where the two dimensional potential is noncentral but separable in polar coordinates r, ϕ. From the above discussion, it is clear that if the potentials in each of the coordinates r and ϕ are chosen to be shape invariant, then the whole problem can be solved algebraically. Calogero's [15] solution of the three-body problem is for the potential

$$V_C=\omega^2/8\sum_{i<j}(x_i-x_j)^2+g\sum_{i<j}(x_i-x_j)^{-2}, \tag{61}$$

where $g>-1/2$ to avoid a collapse of the system. To see why potential given by eq. (61) is solvable, define the Jacobi coordinates

$$R=\tfrac{1}{3}(x_1+x_2+x_3), \tag{62}$$

$$x=\frac{(x_1-x_2)}{\sqrt{2}},\quad y=\frac{(x_1+x_2-2x_3)}{\sqrt{6}}. \tag{63}$$

After elimination of the center of mass part from the Hamiltonian, only the x- and y-degrees of freedom remain, which may be mapped into polar coordinates

$$x=r\sin\phi,\quad y=r\cos\phi. \tag{64}$$

In the above equation,

$$r^2=\tfrac{1}{3}[(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2]. \tag{65}$$

Obviously, the variables r, ϕ have ranges $0 \leq r \leq \infty$ and $0 \leq \phi \leq 2\pi$. It is straightforward to show that

$$(x_1 - x_2) = \sqrt{2}\, r \sin \phi,$$
$$(x_2 - x_3) = \sqrt{2}\, r \sin(\phi + 2\pi/3),$$
$$(x_3 - x_1) = \sqrt{2}\, r \sin(\phi + 4\pi/3). \tag{66}$$

It turns out that V_C is a noncentral but separable potential in the polar coordinates r,ϕ. As a result, the Schrödinger equation separates cleanly in the radial and angular variables so that if the potentials in the radial and angular variables are again SIP then their eigenvalues and eigenfunctions are easily obtained from table 1.

Using this idea, we have recently discovered several new exactly solvable three body problems where one has combined the SIP in the angular variable with the harmonic confinement [17]. The three body scattering problem has also been studied in these cases after dropping the harmonic term [17]. The possibility of replacing the harmonic confinement term with the attractive $1/r$ type interaction has also been considered. Note that in this case one has both discrete and continuous spectra.

7. Isospectral Hamiltonians

In this section, we will describe how one can start from any given one-dimensional potential $V_1(x)$ with atleast one bound state and construct a one parameter family of strictly isospectral potentials [18]. The generalization to obtain an n-parameter family is straight forward [19]. When applied to a reflectionless potential, the n-parameter families provide surprisingly simple expressions for the pure multi-soliton solutions [20] of the Korteweg-de Vries (KdV) and other nonlinear evolution equations. The fact that such families exist has been known for a long time from the inverse scattering approach, but the Gelfand–Levitan approach to finding them is technically much more complicated than the supersymmetry approach described here. Indeed, the advent of SUSY QM has produced a revival of interest in the determination of isospectral potentials.

In the SUSY QM approach, the isospectral families follow when one asks the following question: Suppose $V_2(x)$ is the SUSY partner potential of the original potential $V_1(x)$, and let $W(x)$ be the superpotential such that $V_2(x) = W^2 + W'$ and $V_1(x) = W^2 - W'$. Then, given $V_2(x)$, is the original potential $V_1(x)$ unique i.e., for a given $V_2(x)$, what are the various possible superpotentials $\hat{W}(x)$ and corresponding potentials $\hat{V}_1(x) = \hat{W}^2 - \hat{W}'$? Let us assume [21] that there exists a more general superpotential which satisfies

$$V_2(x) = \hat{W}^2(x) + \hat{W}'(x). \tag{67}$$

Clearly, $\hat{W} = W$ is one of the solutions to eq. (67). To find the most general solution, let us set

$$\hat{W}(x) = W(x) + \phi(x) \tag{68}$$

in eq. (67). We find then that $y(x) = \phi^{-1}(x)$ satisfies the Bernoulli equation

$$y'(x) = 1 + 2Wy \tag{69}$$

whose solution is

$$\frac{1}{y(x)} = \phi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \ln [\mathcal{I}_1(x) + \lambda_1]. \tag{70}$$

Here

$$\mathcal{I}_1(x) \equiv \int_{-\infty}^{x} \psi_1^2(x')\mathrm{d}x'. \tag{71}$$

λ_1 is a constant of integration and $\psi_1(x)$ is the normalized ground state wave function of $V_1(x) = W^2(x) - W'(x)$. Thus the most general $\hat{W}(x)$ satisfying eq. (67) is given by

$$\hat{W}(x) = W(x) + \frac{\mathrm{d}}{\mathrm{d}x} \ln[\mathcal{I}_1(x) + \lambda_1], \tag{72}$$

so that the one parameter family of potentials

$$\hat{V}_1(x) = \hat{W}^2(x) - \hat{W}'(x) = V_1(x) - 2\frac{\mathrm{d}^2}{\mathrm{d}x^2} \ln[\mathcal{I}_1(x) + \lambda_1] \tag{73}$$

have the same SUSY partner $V_2(x)$. One can show that this whole family of potentials are strictly isospectral to $V_1(x)$ (i.e. have same spectrum and same S-matrix) provided that λ_1 does not lie in the interval $-1 \leq \lambda_1 \leq 0$.

To elucidate this discussion, it may be worthwhile to explicitly construct the one-parameter family of strictly isospectral potentials corresponding to the one dimensional harmonic oscillator [18]. In this case

$$W(x) = \frac{\omega}{2} x, \tag{74}$$

so that

$$V_1(x) = \frac{\omega^2}{4} x^2 - \frac{\omega}{2}. \tag{75}$$

The normalized ground state eigenfunction of $V_1(x)$ is

$$\psi_1(x) = \left(\frac{\omega}{2\pi}\right)^{1/4} \exp(-\omega x^2/4). \tag{76}$$

Using eq. (71) it is now straightforward to compute the corresponding $\mathcal{I}_1(x)$. We get

$$\mathcal{I}_1(x) = 1 - \frac{1}{2}\mathrm{erfc}\left(\frac{\sqrt{\omega}}{2}x\right); \quad \mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \mathrm{e}^{-t^2}\mathrm{d}t. \tag{77}$$

In this way one obtains the one parameter family of isospectral potentials and the corresponding wave functions. In figures 1 and 2 we have plotted some of the potentials and the wave functions for the case $\omega = 2$. We see that as λ_1 decreases from ∞ to 0, $\hat{V}_1$ starts developing a minimum which shifts towards $x = -\infty$. Note that as λ_1 finally becomes zero this attractive potential well is lost and we lose a bound state. The remaining potential is called the Pursey potential $V_{\mathrm{P}}(x)$ [22]. An analogous situation occurs in the limit $\lambda_1 = -1$, the remaining potential being the Abraham–Moses potential [23].

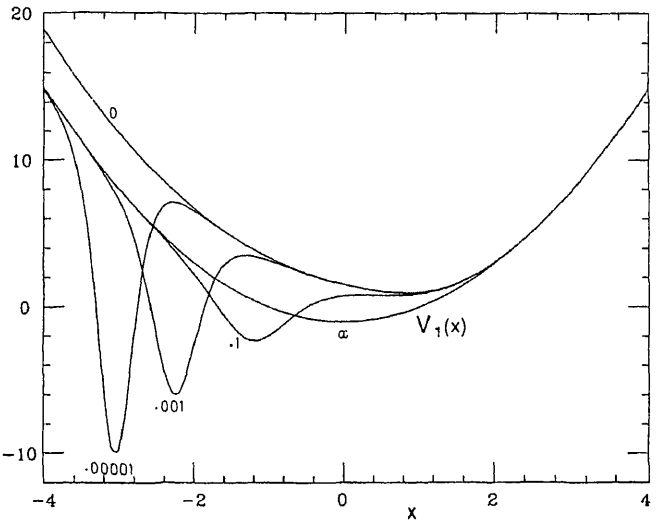


Figure 1. Few members of the family of potentials with energy spectra identical to the one-dimensional harmonic oscillator with $\omega = 2$ and $\hbar = 2m = 1$.

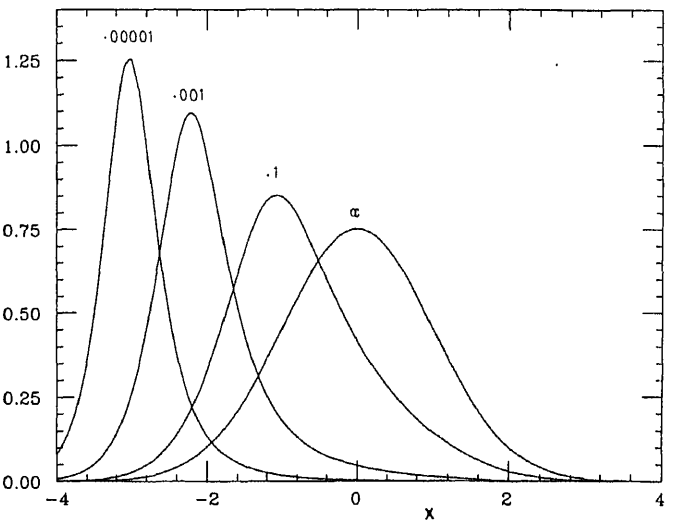


Figure 2. Ground state wave functions for all the potentials shown in figure 1 except the potential with $\lambda_1 = 0$.

8. SUSY and the Dirac equation

There have been many applications of SUSY QM in the context of the Dirac equation. In view of the limitations of space, we shall concentrate here on only one application [24, 4]. In particular, consider the Dirac Lagrangian in $1 + 1$ dimensions with a Lorentz scalar

potential given by

$$\mathcal{L} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - \bar{\psi}\psi\phi. \tag{78}$$

The scalar potential $\phi(x)$ can be looked upon as the static, finite energy, kink solution corresponding to the scalar field Lagrangian

$$\mathcal{L}_{\phi} = \tfrac{1}{2}\partial_{\mu}\phi(x)\partial^{\mu}\phi(x) - V(\phi). \tag{79}$$

Such models have proved quite useful in the context of the phenomenon of fermion number fractionalization which has been seen in certain polymers like polyacetylene. Further, a variant of this model is also relevant in the context of supersymmetric field theories in $1+1$ dimensions. Note that the coupling constant in eqs (78) and (79) has been absorbed in ϕ and $V(\phi)$ respectively.

The Dirac equation following from eq. (78) is

$$i\gamma^{\mu}\partial_{\mu}\psi(x,t) - \phi(x)\psi(x,t) = 0. \tag{80}$$

Let

$$\psi(x,t) = \exp(-i\omega t)\psi(x) \tag{81}$$

so that the Dirac equation reduces to

$$\gamma^0\omega\psi(x) + i\gamma^1\frac{\mathrm{d}\psi(x)}{\mathrm{d}x} - \phi(x)\psi(x) = 0. \tag{82}$$

We choose

$$\gamma^0 = \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^1 = i\sigma^3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \psi(x) = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix} \tag{83}$$

so that we have the coupled equations

$$\begin{aligned} A\psi_1(x) &= \omega\psi_2(x), \\ A^{\dagger}\psi_2(x) &= \omega\psi_1(x), \end{aligned} \tag{84}$$

where

$$A = \frac{\mathrm{d}}{\mathrm{d}x} + \phi(x), \quad A^{\dagger} = -\frac{\mathrm{d}}{\mathrm{d}x} + \phi(x). \tag{85}$$

We can now easily decouple these equations. We get

$$A^{\dagger}A\psi_1 = \omega^2\psi_1, AA^{\dagger}\psi_2 = \omega^2\psi_2. \tag{86}$$

On comparing with the formalism of §2, we see that there is a supersymmetry in the problem and $\phi(x)$ is just the superpotential of the Schrödinger formalism. Further ψ_1 and ψ_2 are the eigenfunctions of the Hamiltonians $H_- \equiv A^+A$ and $H_+ \equiv AA^{\dagger}$ respectively with the corresponding potentials being $V_{\mp}(x) \equiv \phi^2(x) \mp \phi'(x)$. The spectrum of the two Hamiltonians is thus degenerate except that $H_-(H_+)$ has an extra state at zero energy so long as $\phi(x \to \pm\infty)$ have opposite signs and $\phi(x \to +\infty) > 0 (< 0)$.

We thus find a remarkable connection between the scalar Dirac problem and the Schrödinger problem. This connection is very similar to the one between the KdV

equation

$$u_t = 6uu_x - u_{xxx} \tag{87}$$

and the MKdV equation

$$v_t = 6v^2 v_x - v_{xxx}. \tag{88}$$

In particular it is well-known that if v is a solution of the MKdV equation then the two corresponding solutions u_1 and u_2 of the KdV equation are given by the Miura transform i.e.

$$u_1 = v^2 + v_x; \quad u_2 = v^2 - v_x. \tag{89}$$

9. Supersymmetric WKB approximation

The semiclassical WKB method is one of the most useful approximations for computing the energy eigenvalues of the Schrödinger equation. It has a wider range of applicability than standard perturbation theory which is restricted to perturbing potentials with small coupling constants. The purpose of this section is to describe and give applications of the supersymmetric WKB method (henceforth called SWKB) [25,26] which has been inspired by supersymmetric quantum mechanics.

As we have seen in previous sections, for quantum mechanical problems, the main implication of SUSY is that it relates the energy eigenvalues, eigenfunctions and phase shifts of two supersymmetric partner potentials $V_1(x)$ and $V_2(x)$. Combining the ideas of SUSY with the lowest order WKB method, Comtet, Bandrauk and Campbell [25] obtained the lowest order SWKB quantization condition in case SUSY is unbroken and showed that it yields energy eigenvalues which are not only exact for large quantum numbers n (as any WKB approximation scheme should in the semiclassical limit) but which are also exact for the ground state ($n = 0$). We shall now show this in detail.

In lowest order, the WKB quantization condition for the one dimensional potential $V(x)$ is

$$\int_{x_1}^{x_2} \sqrt{2m[E_n - V(x)]}\, dx = (n + 1/2)\hbar\pi, \quad n = 0, 1, 2, \ldots, \tag{90}$$

where x_1 and x_2 are the classical turning points defined by $E_n = V(x_1) = V(x_2)$. For the potential $V_1(x)$ corresponding to the superpotential $W(x)$, the quantization condition (90) takes the form

$$\int_{x_1}^{x_2} \sqrt{2m[E_n^{(1)} - W^2(x) + \frac{\hbar}{\sqrt{2m}} W'(x)]}\, dx = (n + 1/2)\hbar\pi. \tag{91}$$

Let us assume that the superpotential $W(x)$ is formally $O(\hbar^0)$. Then, the W' term is clearly $O(\hbar)$. Therefore, expanding the left hand side in powers of $\hbar$ gives

$$\int_a^b \sqrt{2m[E_n^{(1)} - W^2(x)]}\, dx + \frac{\hbar}{2} \int_a^b \frac{W'(x)dx}{\sqrt{E_n^{(1)} - W^2(x)}} + \cdots = (n + 1/2)\hbar\pi, \tag{92}$$

where a and b are the turning points defined by $E_n^{(1)} = W^2(a) = W^2(b)$. The $O(\hbar)$ term in eq. (92) can be integrated easily to yield

$$\frac{\hbar}{2}\sin^{-1}\left[\frac{W(x)}{\sqrt{E_n^{(1)}}}\right]_a^b. \tag{93}$$

In the case of unbroken SUSY, the superpotential $W(x)$ has opposite signs at the two turning points, that is

$$-W(a) = W(b) = \sqrt{E_n^{(1)}}. \tag{94}$$

For this case, the $O(\hbar)$ term in (93) exactly gives $\hbar\pi/2$, so that to leading order in $\hbar$ the SWKB quantization condition when SUSY is unbroken is [25, 26]

$$\int_a^b \sqrt{2m[E_n^{(1)} - W^2(x)]}\,\mathrm{d}x = n\hbar\pi, \quad n = 0, 1, 2, \ldots . \tag{95}$$

Proceeding in the same way, the SWKB quantization condition for the potential $V_2(x)$ turns out to be

$$\int_a^b \sqrt{2m[E_n^{(2)} - W^2(x)]}\,\mathrm{d}x = (n+1)\hbar\pi, \quad n = 0, 1, 2, \ldots . \tag{96}$$

Some remarks are in order at this stage.

(i) For $n = 0$, the turning points a and b in eq. (95) are coincident and $E_0^{(1)} = 0$. Hence SWKB is exact by construction for the ground state energy of the Hamiltonian $H_1 = (-\hbar^2/2m)\mathrm{d}^2/\mathrm{d}x^2 + V_1(x)$.

(ii) On comparing eqs (95) and (96), it follows that the lowest order SWKB quantization condition preserves the SUSY level degeneracy i.e. the approximate energy eigenvalues computed from the SWKB quantization conditions for $V_1(x)$ and $V_2(x)$ satisfy the exact degeneracy relation $E_{n+1}^{(1)} = E_n^{(2)}$.

(iii) Since the lowest order SWKB approximation is not only exact as expected for large n, but is also exact by construction for $n = 0$, hence, unlike the ordinary WKB approach, the SWKB eigenvalues are constrained to be accurate at both ends of allowed values of n at least when the spectrum is purely discrete. One can thus reasonably expect better results than the WKB scheme. This is born out by several calculations.

(iv) It has been proved that the lowest order SWKB approximation reproduces the exact bound state spectrum of any SIP as given in table 1 [27]. This is in contrast to the lowest order WKB which is exact for only one potential i.e. the one dimensional harmonic oscillator.

(v) Further it has been explicitly checked for all SIPs of table 1 that up to $O(\hbar^6)$, there are indeed no corrections to the lowest order exact result [28]. These results can be extended to all orders in $\hbar$.

(vi) It has recently been shown [29] that the lowest order SWKB is not exact for the newly discovered class [5, 6] of SIPs for which the parameters a_2 and a_1 are related by scaling $a_2 = qa_1$.

(vii) Recently, quantization condition has also been obtained in the case of broken SUSY [30, 32] and it has been proved that even here the lowest order condition gives exact spectra in the case of the three SIP with broken SUSY [31] and all higher order corrections are zero [30, 32, 33].

10. Bound states in the continuum

In 1929, von Neumann and Wigner [34] realized that it was possible to construct potentials which have quantum mechanical bound states embedded in the classical energy continuum (BICs). Further developments, by many authors have produced more examples and a better understanding of the kind of potential that can have such bound states, although there is not as yet a fully systematic approach. Capasso *et al* [35] have recently reported direct evidence for BICs by constructing suitable potentials using semiconductor heterostructures grown by molecular beam epitaxy. Finally, it is interesting to note that BICs have found their way into a recently written text [36].

In this section, we show how one can start from a potential with a continuum of energy eigenstates, and use the methods of SUSY QM to generate families of potentials with bound states in the continuum [BICs] [37]. Basically, one is using the technique of generating isospectral potentials (discussed in § 7) but this time starting from states in the continuum. The method preserves the spectrum of the original potential except it adds these discrete BICs at selected energies.

Consider any spherically symmetric potential $V(r)$ which vanishes as $r \to \infty$. The radial s-wave Schrödinger equation for the reduced wave function $u(r)$ (in units where $\hbar = 2m = 1$) is

$$-u'' + V(r)u(r) = E\,u(r), \tag{97}$$

where we have scaled the energy and radial variables such that all quantities are dimensionless. Equation (97) has a classical continuum of positive energy solutions which are clearly not normalizable.

As we have seen in § 7, the isospectral potential $\hat{V}(r;\lambda)$ is given in terms of the original potential $V(r)$ and the original ground state wave function $u_0(r)$ by [21, 18, 19]

$$\hat{V}(r;\lambda) = V(r) - 2[\ln(I_0+\lambda)]'' = V(r) - \frac{4u_0u_0'}{I_0+\lambda} + \frac{2u_0^4}{(I_0+\lambda)^2}, \tag{98}$$

where

$$I_0(r) \equiv \int_0^r u_0^2(r')\mathrm{d}r'. \tag{99}$$

Except in this section, u_0 was taken to be the nodeless, normalizable ground state wave function of the starting potential $V(r)$. However, it is easy to generalize the above equations to the case where $u_0(r)$ is any solution of eq. (97) with arbitrary energy E_0. Our results are best summarized in the following statement.

Theorem. *Let $u_0(r)$ and $u_1(r)$ be any two nonsingular solutions of the Schrödinger equation for the potential $V(r)$ corresponding to arbitrarily selected energies E_0 and E_1*

respectively. Construct a new potential $\hat{V}(r;\lambda)$ as prescribed by eq. (98). Then, the two functions

$$\hat{u}_0(r;\lambda) = \frac{u_0(r)}{I_0 + \lambda}, \tag{100}$$

and

$$\hat{u}_1(r;\lambda) = (E_1 - E_0)u_1 + \hat{u}_0\, W(u_0, u_1), \tag{101}$$

[where W denotes the Wronskian, $W(u_0, u_1) \equiv u_0 u_1' - u_1 u_0'$] are solutions of the Schrödinger equation for the new potential $\hat{V}(r;\lambda)$ corresponding to the same energies E_0 and E_1.

Example. Free particle on the half line. Consider a free particle on the half line ($V \equiv 0$ for $0 \leq r < \infty$). We choose $u_0 = \sin kr$, the spherical wave solution, corresponding to energy $E_0 = k^2 > 0$, which vanishes at $r = 0$. The integral I_0 given in eq. (99) becomes

$$I_0 = [2kr - \sin(2kr)]/(4k). \tag{102}$$

We observe that $I_0 \to r/2$ as $r \to \infty$.

The potential family $\hat{V}$, defined in eq. (98) becomes

$$\hat{V}(r;\lambda) = \frac{32k^2 \sin^4 kr}{D_0^2} - \frac{8k^2 \sin(2kr)}{D_0} \tag{103}$$

with

$$D_0(r;\lambda) = 2kr - \sin(2kr) + 4k\lambda. \tag{104}$$

$\hat{V}$ has a BIC at energy $E_0 = k^2$ with wave function

$$\hat{u}_0(\lambda) = 4k \sin kr / D_0. \tag{105}$$

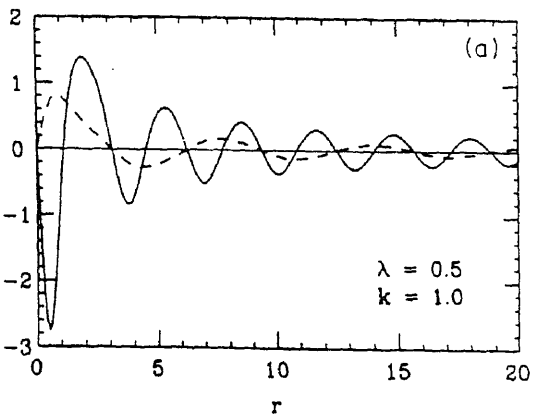


Figure 3. Potential $\hat{V}(r)$ (solid lines) possessing one bound state in the continuum (BIC) obtained by starting from a free particle in the case $\lambda = 0.5$ and $E = 1$. The BIC wave function is shown by the dashed line.

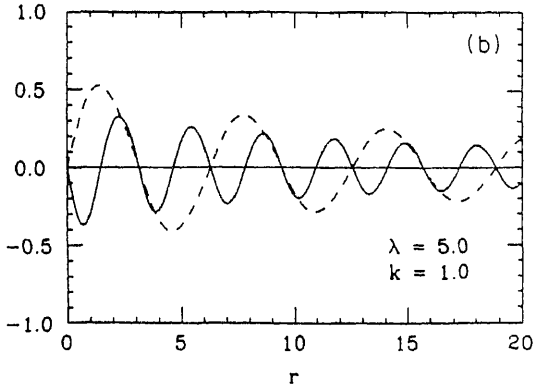


Figure 4. Potential $\hat{V}(r)$ in the case $\lambda = 0.5$ and $E = 1$. The BIC wave function is again shown by a dashed line.

For special values of the parameters k and λ, the potential $\hat{V}$ and its BIC wave functions are shown in figures 3 and 4. The original null potential has now become an oscillatory potential which asymptotically has a $1/r$ envelope. The new wave function at $E_0 = k^2$ also has an additional damping factor of $1/r$ which makes it square integrable. The parameter λ which appears in the denominator function $D_0(r;\lambda)$ plays the role of a damping distance; its magnitude indicates the value of r at which the monotonically growing integral I_0 becomes a significant damping factor, both for the new potential and for the new wave function. The parameter λ must be restricted to values greater than zero in order to avoid infinities in $\hat{V}$ and in the wave functions. In the limit $\lambda \to \infty$, $\hat{V}$ becomes identical to V.

Acknowledgments

It is a pleasure to thank Profs A C Das, J C Parikh and S B Khadkikar for organizing an excellent meeting "Fundamentals of Physics and Astrophysics" during 14–16 March, 1995 on the occasion of the sixtieth birthday of Prof. R K Verma and for the warm and cordial hospitality.

References

[1] See for example, *Supersymmetry in Physics* edited by V A Kostelecky and D K Campbell (North Holland, 1985)
[2] E Witten, *Nucl. Phys.* **B188**, 513 (1981)
[3] F Cooper and B Freedman, *Ann. Phys.* **146**, 262 (1983)
[4] F Cooper, A Khare and U P Sukhatme, *Phys. Rep.* **251**, 267 (1995) and references therein
[5] A Khare and U Sukhatme, *J. Phys.* **A26**, L901 (1993)
[6] D Barclay, R Dutt, A Gangopadhyaya, A Khare, A Pagnamenta and U Sukhatme, *Phys. Rev.* **A48**, 2786 (1993)
[7] L Infeld and T E Hull, *Rev. Mod. Phys.* **23**, 21 (1951)
[8] L Gendenshtein, *JETP Lett.* **38**, 356 (1983)

[9] F Cooper, J N Ginocchio and A Khare, *Phys. Rev.* **D36**, 2458 (1987)
[10] R Dutt, A Khare and U Sukhatme, *Am. J. Phys.* **56**, 163 (1988)
[11] G Levai, *J. Phys.* **A22**, 689 (1989)
[12] J Dabrowska, A Khare and U Sukhatme, *J. Phys.* **A21**, L195 (1988)
[13] A Khare and U Sukhatme, *J. Phys.* **A21**, L501 (1988)
[14] A Khare and R K Bhaduri, *Am. J. Phys.* **62**, 1008 (1994)
[15] F Calogero, *J. Math. Phys.* **10**, 2191 (1969)
Also see, B Sutherland, *J. Math. Phys.* **12**, 246 (1971)
[16] S B Isakov, *Int. J. Mod. Phys.* **A9**, 2563 (1994)
Also see A P Polychronakos, *Nucl. Phys.* **B324**, 597 (1989)
[17] A Khare and R K Bhaduri, *J. Phys.* **A27**, (1994) (in press)
[18] A Khare and U Sukhatme, *J. Phys.* **A22**, 2847 (1989)
[19] W-Y Keung, U Sukhatme, Q Wang and T Imbo, *J. Phys.* **A22**, 987 (1989)
[20] Q Wang, U Sukhatme, W-Y Keung and T Imbo, *Mod. Phys. Lett.* **A5**, 525 (1990)
[21] M M Nieto, *Phys. Lett.* **B145**, 208 (1984)
[22] D Pursey, *Phys. Rev.* **D33**, 1098, 1103, 2267 (1986)
M Luban and D Pursey, *Phys. Rev.* **D33**, 431 (1986)
[23] P Abraham and H Moses, *Phys. Rev.* **A22**, 1333 (1980)
[24] B Thaller, *The Dirac Equation* (Springer Verlag, 1992)
[25] A Comtet, A Bandrauk and D K Campbell, *Phys. Lett.* **B150**, 159 (1985)
[26] A Khare, *Phys. Lett.* **B161**, 131 (1985)
[27] R Dutt, A Khare and U Sukhatme, *Phys. Lett.* **B181**, 295 (1986)
[28] R Adhikari, R Dutt, A Khare and U Sukhatme, *Phys. Rev.* **A38**, 1679 (1986)
[29] D Barclay, A Khare and U Sukhatme, *Phys. Lett.* **A183**, 263 (1993)
[30] A Inomata and G Junker, *Lectures on Path Integration* edited by H Cerdeira, S Lundqvist, D Mugnai, A Ranfangi, Y Sa-yakanit and L Schulman (World Scientific, Singapore, 1992); *Symposium on Adv. Topics in Quantum Mechanics* edited by J Q Liang *et al* (Science Press, Beijing, 1992)
[31] R Dutt, A Gangopadhyaya, A Khare, A Pagnamenta and U Sukhatme, *Phys. Lett.* **A174**, 363 (1993)
[32] R Dutt, A Gangopadhyaya, A Khare, A Pagnamenta and U Sukhatme, *Phys. Rev.* **A48**, 1845 (1993)
[33] N R Murali, T R Govindarajan and A Khare, IMSc 94/48, IP-BBSR/94-74; submitted to *Phys. Rev.* **A52**, 4259 (1995)
[34] J von Neumann and E Wigner, *Z. Phys.* **30**, 465 (1929)
[35] F Capasso, C Sirtori, J Faist, D L Sirco, S Chu and A Cho, *Nature* **358**, 5656 (1992)
[36] L E Ballentine, *Quantum Mechanics* (Prentice Hall, 1990) pp. 205
[37] J Pappademos, U Sukhatme and A Pagnamenta, *Phys. Rev.* **A48**, 3525 (1993)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 65–69

Violation of signal locality and unitarity in a merger of quantum mechanics and general relativity

PARTHA GHOSE
S.N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake, Calcutta 700 091, India

Abstract. It is shown that a violation of signal locality and unitarity occur in a particular merger of quantum mechanics and general relativity.

Keywords. Signal locality; quantum theory; quasiclassical gravity.

PACS Nos 03.65; 04.00

In spite of much effort that has been spent over decades, it has not been possible to formulate a completely satisfactory quantum theory of gravity. This has led some to suspect that quantum theory and the general theory of relativity are incompatible in their present forms [1]. In the absence of a satisfactory quantum theory of gravity, it has been traditional to use a semi-classical theory in which the metric of space-time is determined by the local expectation value of the stress-energy-momentum tensor of quantized matter via the classical field equations of Einstein:

$$G_{\mu\nu} = 8\pi\langle T_{\mu\nu}(x)\rangle. \tag{1}$$

Although such a semi-classical theory is expected to be reliable far away from the Planck scale, Kibble has emphasized that it implies a non-linear modification of quantum theory [2]. He has also suggested a Gedanken experiment which shows that its predictions are inconsistent with those of quantum theory which should automatically incorporate gravity if one takes the view that the universe is wholly quantum and that classical Einsteinian gravity is only an approximate description of phenomena that emerge in certain situations. Let me describe a slight variation of Kibble's Gedanken experiment. Imagine a device like a Stern–Gerlach apparatus that splits the wavefunction ψ of an incoming particle (whose mass M is sufficiently large for the purpose in hand) into a superposition of two orthogonal pieces that are eventually registered by one of two detectors D_1 and D_2 localized at x_1 and x_2. The final wavefunction of the combined system (particle plus the detectors) is then

$$\Phi = \frac{1}{\sqrt{2}}[\psi_1(x)D_1(x) + \psi_2(x)D_2(x)]. \tag{2}$$

A small test mass m is placed midway between the two detectors. Every time a detector fires, m is attracted towards it according to quantum theory. According to traditional

semi-classical theory, however, the metric due to the particle wavefunction ψ would be determined by

$$\tfrac{1}{2}[\langle T_{\mu\nu}(x_1)\rangle + \langle T_{\mu\nu}(x_2)\rangle] \tag{3}$$

which does not change when a detector fires, and so m does not move at all.

Recently a proposal has been made [3] for a quasiclassical theory of gravity in which one has to solve the coupled equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu}(\psi), \quad T^{\mu\nu}_{;\nu} = 0, \tag{4}$$

$$i\hbar\partial_{;t}\psi = H(-i\hbar\partial_{;x})\psi, \tag{5}$$

where the metric is supposed to be quasiclassical and the matter part first quantized. The key feature is that the classical variables become correlated to the state of the quantum variables. If the quantum state is a superposition of different states, the classical variables need not have a definite value but may take a distribution of values depending on the quantum state. Effectively, the quantum variables act like a source of noise on the classical evolution. These variables are then not entirely classical in nature but may more properly be called *quasiclassical*. In contrast to the traditional semi-classical approach (eq. (1)), this approach maintains the linearity of quantum mechanics through the coupled equations (4) and (5) while allowing the values of the quasiclassical variables to depend on the quantum state. This provides a means of computing the backreaction that quantum variables have on the evolution of classical variables without having to make a full semi-classical analysis. If this approach extends to quantum field theory, it has potentially far reaching consequences. By keeping the gravitational field quasiclassical, it reopens the debate on the necessity to quantize gravity [4]. Alternatively, when viewed as an approximation to a fully quantized theory of gravity far away from the Planck scale, it improves on conventional quantum field theory in curved space-time by taking into account the backreaction of quantized matter fields on the classical space-time background. Such effects are expected to be crucial for a complete understanding of information loss in black hole evaporation [5].

The consistency of this approach with linear quantum mechanics can be checked in the case of Kibble's Gedanken experiment discussed above. Since the metric due to the matter wavefunction Φ is determined in this approach by

$$\tfrac{1}{2}[T_{\mu\nu}(x_1)|D_1|^2 + T_{\mu\nu}(x_2)|D_2|^2] \tag{6}$$

rather than by (3), each time a detector fires, Φ 'collapses' to only one of the terms in (2), and so the mass m moves. This is a general consequence of the formalism of quantum mechanics and is independent of interpretations. For example, even the many-worlds interpretation in which there is no collapse in the universe as a whole would imply this for every branch into which the universe splits. This means that the metric changes each time a detector fires, and m is attracted towards it, which is consistent with quantum theory.

The question arises as to whether consistency with linear quantum mechanics can be demonstrated most generally. I will now show that there is at least one case in which an inconsistency arises. This happens because of collapse-induced non-locality. Although quantum mechanics implies Bell-type non-locality through non-local quantum correlations that cannot be reproduced in any local hidden variable theory, it is consistent with

signal locality, i.e., the prohibition of superluminal signals at the statistical level, even though the notion of wavefunction collapse as an actual physical transition from a pure to a mixed state implies non-local influences at the level of individual events [6]. In a realist interpretation of these events such non-local influences constitute a definite violation of the principle of relativity.

Consider the two-particle position and momentum correlated entangled state

$$|\psi\rangle = \frac{1}{\sqrt{2}}[\,|A\rangle|B\rangle\delta^{(3)}(x_A - x_B - d(t)) + |A'\rangle|B'\rangle\delta^{(3)}(x_{A'} - x_{B'} - d(t))], \qquad (7)$$

where $|A\rangle$ and $|A'\rangle$ are two possible orthogonal position eigenstates of particle A moving to the left (say) of the source, and $|B\rangle$ and $|B'\rangle$ are the corresponding correlated position eigenstates of particle B moving to the right; $d(t)$ is the distance between the particles at the time t determined by their relative velocity. The density operator $\rho = |\psi\rangle\langle\psi|$ is non-diagonal, and one can calculate the reduced density operator of one of the particles (A or B) by tracing over the states of the other. For example,

$$\rho_B = \mathrm{Tr}_A\rho = |B\rangle\langle B| + |B'\rangle\langle B'|, \quad \mathrm{Tr}\rho_B = 1. \qquad (8)$$

It is the orthogonality of the two possible states $|A\rangle$ and $|A'\rangle$ that results in ρ_B being diagonal. This implies that effectively ρ_B describes an incoherent mixture of $|B\rangle$ and $|B'\rangle$, and that these do not interfere with each other even though the two-particle state $|\psi\rangle$ remains entangled. This is because when the system is prepared in the state $|\psi\rangle$, one could determine which path particle B takes by placing detectors in the beams A and A', and this determination does not disturb particle B in any mechanical sense, as emphasized by Einstein, Podolsky and Rosen.

According to Feynman [7], 'if an experiment is performed which is capable of determining whether one or another alternative is actually taken, the probability of the event is the sum of the probabilities for each alternative. The interference is lost'. This rule can be extended in the spirit of Einstein, Podolsky and Rosen by replacing 'is performed' in Feynman's rule by 'could be performed without disturbing the system'. This is a novel feature of two-particle interferometry [8]. It has the important corollary that a localized measurement on A cannot affect ρ_B (the trace over the possible states of A having already been taken in computing it), and hence the expectation value of all local, hermitean operators $\hat{O}$ in the Hilbert space of B, given by

$$\langle\hat{O}\rangle = \mathrm{Tr}\,\rho_B\hat{O}, \qquad (9)$$

remains unchanged when such a measurement is made. This guarantees that no superluminal signal can be sent utilizing entangled states and collapse-induced non-locality.

The situation is different if gravitational effects are included in the way suggested in [3]. When a localized measurement is made on A by placing detectors at x_A and $x_{A'}$, the state of the combined system of the two particles A and B and the detectors can be written as

$$|\Phi\rangle = \frac{1}{\sqrt{2}}[|A\rangle|B\rangle|D_A\rangle\delta^{(3)}(x_A - x_B - d(t)) + |A'\rangle|B'\rangle|D_{A'}\rangle\delta^{(3)}(x_{A'} - x_{B'} - d(t))], \qquad (10)$$

where $|D_A\rangle$ and $|D_{A'}\rangle$ are the states of the two detectors coupled to $|A\rangle$ at x_A and $|A'\rangle$ at $x_{A'}$ respectively. On completion of the measurement, 'collapse' or state vector reduction occurs, i.e., only one of the two terms in (10) survives. If it is the first term, the particle B is instantaneously projected to its state $|B\rangle$ localized at x_B; if it is the second term, it is projected to its state $|B'\rangle$ localized at $x_{B'}$. This localization of B is represented by a change of its wavefunction from the entangled ψ to $\phi_B(x) = \langle x|B\rangle = \delta^{(3)}(x - x_B)$ or $\phi_{B'}(x) = \langle x|B'\rangle = \delta^{(3)}(x - x_{B'})$. The local stress–energy–momentum density $T_{\mu\nu}(x)$ of B (which is not an observable) must therefore change from

$$\tfrac{1}{2}[\psi^*(x)\hat{T}_{\mu\nu}\psi_B + \psi^*_{B'}(x)\hat{T}_{\mu\nu}\psi_{B'}(x)] \tag{11}$$

before the measurement to $\phi^*_B\hat{T}_{\mu\nu}\phi_B(x)$ or $\phi^*_{B'}(x)\hat{T}_{\mu\nu}\phi_{B'}(x)$ after the measurement, resulting in an instantaneous and non-local change of the metric field around x_B or $x_{B'}$ to a Schwarzchild form via eq. (4) [9]. Although such a change may be quantitatively insignificant and for all practical purposes ignored in many cases, in a theory in which the gravitational field is regarded as being fundamentally classical or quasiclassical, it would already imply a *non-local* and *real factual*, and therefore in principle observable, change at the level of individual events that contradicts relativity theory.

Even in the alternative scenario in which the quasiclassical gravitational field is regarded as an approximation to a fully quantum theory of gravity, this non-local change of metric due to 'collapse' has an unacceptable consequence. 'Collapse' or state vector reduction is a non-causal, non-linear and non-unitary change during which the normal dynamical law of evolution (linear and unitary (eq. (5)) remains suspended. Once it has occurred, normal dynamical evolution resumes but with a difference in the presence of gravity—with the covariant derivatives in eq. (5) changed non-causally by the 'collapse'–induced metric change. Consequently, the new states $|\mathcal{B}\rangle$ and $|\mathcal{B}'\rangle$ that are solutions to eq. (5) with the changed metric must be related to the old ones $|B\rangle$ and $|B'\rangle$ by a non-unitary operator C, i.e., $|\mathcal{B}\rangle = C|B\rangle$, $|\mathcal{B}'\rangle = C|B'\rangle$ with $C^\dagger \neq C^{-1}$. I must emphasize that C is *not* to be confused with the projection operator for collapse. Then the new density operator for B has the property

$$\mathrm{Tr}\,\rho'_B = \mathrm{Tr}(|\mathcal{B}\rangle\langle\mathcal{B}| + |\mathcal{B}'\rangle\langle\mathcal{B}'|) \tag{12}$$

$$= \mathrm{Tr}\,C\rho_B C^\dagger \neq 1. \tag{13}$$

Consequently, the expectation values of all observables change, unitarity is violated and superluminal signalling becomes possible. It is clearly only gravity (as formulated in [3]) that is capable of exposing this unacceptable and ugly head of 'collapse'. This inconsistency, if true, can only be removed by changing standard measurement theory to make it consistent with relativistic invariance. Although 'retarded collapse' (the effect of collapse propagating within the light-cone) looks like an attractive possibility in this context, it is incompatible with quantum mechanics because it is equivalent to a local model and therefore compatible with Bell's inequalities which are violated by quantum mechanics.

This violation of signal locality, a fundamental tenet of relativity theory, is a feature of the proposal in [3] that is, however, not shared by the traditional semi-classical approach because of the use of the local expectation value of the stress-energy-momentum tensor in eq. (1), which is an observable and does not change with measurements. But, as Kibble

has shown through his Gedanken experiment, this very feature makes it inconsistent with quantum theory.

All this tends to confirm the suspicion that the two pillars of modern physics, quantum theory and general relativity, are probably incompatible in their present forms.

Acknowledgement

The lecture was based on work done in collaboration with A S Majumdar [10]. We are grateful to Dipankar Home for many helpful discussions. I feel deeply honoured at being able to offer this essay as a tribute to Prof. R K Varma on his sixtieth birthday. The work was done as part of a reseach project sanctioned by the Department of Science and Technology, Government of India.

References

[1] R Penrose, *Newton, quantum theory and reality* in *300 years of gravitation* edited by S W Hawking and W Israel (Cambridge University Press, Cambridge, 1987)
[2] T W B Kibble, in *Quantum Gravity 2* edited by C J Isham, R Penrose and D W Sciama (Clarendon Press, Oxford, 1981)
[3] A Anderson, *Phys. Rev. Lett.* **74**, 621 (1995)
[4] R P Feynman, *Lectures on Gravitation* (Caltech, 1962–63) unpublished
[5] See, for example, D Page, in *Proceedings of the 5th Canadian Conference on General Relativity and Relativistic Astrophysics* edited by R B Mann and R G McLenaghan (World Scientific, Singapore, 1994)
[6] D Home and R Nair, *Phys. Lett.* **A187**, 224 (1994)
[7] R P Feynman, R B Leighton and M Sands, *Lectures on Physics* (Addison-Wesley, Reading, MA, 1963) vol. 1
[8] G Jaegger, M A Horne and A Shimony, *Phys. Rev.* **A48**, 1023 (1993)
[9] D V Ahluwalia, *Phys. Lett.* **B339**, 301 (1994)
[10] P Ghose and A S Majumdar, SNBNCBS preprint (1995)

—journal of
physics

July 1997
pp. 71–80

Quark confinement in $2+1$ dimensional pure Yang–Mills theory

SPENTA R WADIA
Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

Abstract. We report some progress on the quark confinement problem in $2+1$ dim. pure Yang–Mills theory, using Euclidean instanton methods. The instantons are regularized Wu–Yang 'monopoles', whose long range Coulomb field is screened by collective effects. Such configurations are stable to small perturbations unlike the case of singular, undressed monopoles. Using exact non-perturbative results for the 3-dim. Coulomb gas, where Debye screening holds for arbitrarily low temperatures, we show in a self-consistent way that a mass gap is dynamically generated in the gauge theory. The mass gap also determines the size of the monopoles. We also identify the disorder operator of the model in terms of the Sine–Gordon field of the Coulomb gas and hence obtain a dual representation whose symmetry is the centre of $SU(2)$.

Keywords. Quark; confinement; instanton; disorder.

PACS No. 11.15

1. Introduction

In this talk we will discuss some progress towards an understanding of the dynamics of pure Yang–Mills theory in $2+1$ dims. [1]. Our discussion uses the Euclidean instanton approach which has been successful in elucidating the dynamics of $2+1$ dim. Yang–Mills theory in the presence of adjoint Higgs [2, 3]. However the gauge theory without the Higgs fields is a much more difficult problem for several reasons. Firstly, perturbation theory, though ultraviolet finite, is hopelessly infrared divergent. Secondly, because this theory has only one length scale set by the gauge coupling g, unlike the theory with the Higgs fields, one does not have the luxury of having another length scale m_W to enable a controlled semi-classical approximation. The classical monopole configurations are *singular* configurations in the continuum limit and hence the renormalization of the monopole gas is very nontrivial.

It is possible to regulate the singularity by modifying the fields inside a 'core' of some size λ. The classical action now depends on λ. We do not vary the size λ of the monopoles. Rather we treat it as a parameter of the theory to be self consistently determined by the mass gap.

In the absence of Higgs fields a simple monopole is unstable to small fluctuations of gauge fields around it. The instability can be traced to the long range Coulomb tail of the monopole in the small oscillations operator. However a Coulomb gas of magnetic monopoles in 3 dims. always screens the long range magnetic field of a monopole. Therefore the small oscillations of the gauge field around a screening plasma of

monopoles need not be unstable. In this way we construct a self-consistent picture of instanton gas. The Debye length then self-consistently determines the size of the monopoles. The resulting long distance effective theory we obtain is a sine-Gordon theory in $2+1$ dim. We also show that the Sine–Gordon theory is actually a dual representation of $2+1$ YM theory in the sense that the disorder (magnetic) operators of this theory have a local formula in terms of the Sine–Gordon field χ: $\phi_D = e^{(i/2)\chi(x)}$.

2. Topology of the gauge condition

We begin with a review the topological properties of the vacuum of $SU(2)$ Yang–Mills theory in $2+1$ dimensions following [4]. The physical degrees of freedom are most transparent in a unitary type of gauge. For pure gauge theories this is defined as follows. Consider some local operator $X(x)$ which transforms according to the adjoint representation of $SU(2)$. The unitary gauge is now defined by

$$[X(x), \tau_3] = 0, \tag{1}$$

(τ^i, $i = 1,2,3$ stands for the Pauli matrices). This gauge condition retains the $U(1)$ generated by τ^3 as an unbroken symmetry. If the model contains adjoint Higgs fields in addition to the gluons, X may be the Higgs field itself and (1) is the conventional unitary gauge in such a Georgi–Glashow model. In pure Yang–Mills theory X has to be constructed out of the gauge fields alone.

We may write the matrix operator X in the form

$$X = \lambda I + \sum_{a=1}^{3} \epsilon_a(x)\tau^a, \tag{2}$$

where I is the identity matrix. Then the points x_0 where

$$\epsilon^a(x_0) = 0 \tag{3}$$

are singularities of the gauge condition (1). As shown in [4] these singularities are nothing but magnetic monopoles with respect to the unbroken $U(1)$. In $3+1$ dimensions these monopoles are 'particles' which represent dynamical degrees of freedom (in this gauge) other than the conventional fields. In $2+1$ dimensions these monopoles are instanton configurations which will play a crucial role in determining the properties of the vacuum.

In the Georgi–Glashow model the monopoles are given by the well-known 't Hooft–Polyakov solution, while for pure gauge theory they are described by the regularized Wu–Yang solution.

3. Confinement in pure Yang–Mills theory

In pure Yang–Mills theory as we have mentioned there is no Higgs field. This has two consequences. First, in the absence of the second scale (i.e. m_W), the perturbation expansion is infrared divergent and the theory cannot be defined perturbatively in the infinite volume limit. Secondly the classical monopole solutions are Wu–Yang

expansion around a plasma of such monopoles and, as explained in the introduction, argue that these monopoles are stable against fluctuations.

3.1 *The monopole solution*

We will consider monopoles which have some size λ. The field due to single monopole at $\mathbf{x} = 0$ is given in the unitary gauge by

$$\begin{aligned}
\tilde{A}_\mu(x) &= \frac{1}{2}\begin{bmatrix} q\tilde{A}^3_\mu & \tilde{A}^1_\mu - iq\tilde{A}^2_\mu \\ \tilde{A}^1_\mu + iq\tilde{A}^2_\mu & -q\tilde{A}^3_\mu \end{bmatrix}, \\
\tilde{A}^1_\mu(x) &= -\frac{K(r/\lambda)}{r}\,[\hat{\phi}\cos\phi + \hat{\theta}\sin\phi]_\mu, \\
\tilde{A}^2_\mu(x) &= \frac{K(r/\lambda)}{r}\,[-\hat{\phi}\sin\phi + \hat{\theta}\cos\phi]_\mu, \\
\tilde{A}^3_\mu(x) &= -\frac{1}{r}\tan\frac{\theta}{2}\,[\hat{\phi}]_\mu = D_\mu
\end{aligned} \tag{4}$$

in spherical coordinates. Here $K(r/\lambda)$ is a structure function regulating the fields at $r = 0$. The function $K(r/\lambda)$ goes to 1 as $r \to 0$ as follows

$$K(r/\lambda) \sim 1 - \frac{r^2}{\lambda^2} \quad \text{for } r \to 0 \tag{5}$$

while at $r = \lambda$, $K(r/\lambda) = K'(r/\lambda) = 0$ and remains zero for $r > \lambda$. Furthermore $K'(r/\lambda)$ is continuous at $r = \lambda$. As shown by Banks, Myerson and Kogut [6], one can choose a $K(r/\lambda)$ so that the configuration (4) is a classical solution. The action of single monopole is

$$s = \frac{4\pi}{g^2}\int_0^\infty \mathrm{d}r\left[\left(\frac{\mathrm{d}K}{\mathrm{d}r}\right)^2 + \left(\frac{K^2-1}{2r^2}\right)\right]. \tag{6}$$

Note that the monopole field is abelian outside the monopole core. Consider a gas of such monopoles such that if the positions of the monopoles are x_a, x_b etc., one always has $|x_a - x_b| > 2\lambda$. The field configuration in the regions outside the core of the monopoles is approximated by

$$\mathcal{A}^{\mathrm{cl}}_\mu(x) = \sum_{a=1}^{N} \tilde{A}_\mu(x - x_a), \tag{7}$$

where x_a denotes a monopole position and N is the number of monopoles. In our self-consistent approach, we assume that (7) represents the dominant field configuration in the Euclidean path integral. The total action of this gas is

$$S_{\mathrm{cl}} = Ns + \frac{2\pi}{g^2}\sum_{a\neq b}\frac{q_a q_b}{|x_a - x_b|}. \tag{8}$$

Finally we record the single monopole field configuration in the radial gauge

$$A_\mu^a = -\frac{\epsilon_{\mu aj}x^j}{r^2}(1 - K(r/\lambda)). \tag{9}$$

3.2 *The path integral*

The monopole configuration is used to evaluate the path integral by the saddle point method. One expands the field around $\mathcal{A}_\mu^{\rm cl}$ which is a classical solution outside the core. The form of the solution inside the core is unimportant for our purposes:

$$\mathcal{A}_\mu = \mathcal{A}_\mu^{\rm cl} + g a_\mu. \tag{10}$$

The path integral may be formally written as

$$Z = \int \prod_x \mathrm{d}\mathcal{A}_\mu(x)$$
$$\exp(-S_{\rm cl}) \exp\left(-\int a D^2 a \mathrm{d}^3x\right), \tag{11}$$

where $S_{\rm cl}$ is given by (8). Here D^2 is the stability operator for small fluctuations:

$$\int a D^2 a \, \mathrm{d}^3x \equiv \int \mathcal{L}''(\mathcal{A}_{\rm cl}) a^2 \mathrm{d}^3x, \tag{12}$$

where $\mathcal{L}''(\mathcal{A}_{\rm cl})$ is the second functional derivative of the Lagrangian density evaluated at $\mathcal{A}_{\rm cl}$.

Equation (11) as it stands is meaningless since D^2 has a zero eigenvalue for each symmetry of the original Lagrangian. The local gauge symmetry is fixed by requiring the fluctuations to satisfy the background gauge condition

$$\nabla_\mu(\mathcal{A}^{\rm cl}) a_\mu(x) = 0, \tag{13}$$

where $\nabla_\mu(\mathcal{A}^{\rm cl})$ is the covariant derivative evaluated at the configuration $\mathcal{A}^{\rm cl}$:

$$\nabla_\mu(\mathcal{A}_{\rm cl}) \equiv \partial_\mu + i[\mathcal{A}_\mu^{\rm cl}, \,]. \tag{14}$$

This gives rise to the usual Fadeev-Popov determinant ($\det \Delta_{\rm FP}$). The zero modes arising from the breaking of the global translation invariance and the $U(1)$ transformations obeying the background gauge condition but nonvanishing at infinity are replaced by integration over corresponding collective coordinates. Finally we have to sum over all N with the standard division by $N!$. One finally has the formal expressions

$$Z = \sum_{N=0}^{\infty} \frac{1}{N!} Q_N$$

$$Q_N = \sum_{\{q_a\}} \left(\frac{8}{\sqrt{\pi}} s^{3/2}\right)^N \int \prod_{a=1}^{N} \mathrm{d}x_a \mathrm{e}^{-S_{\rm cl}} \mathcal{J} \tag{15}$$

$$\mathcal{J} = \left(\frac{\det D^2}{\det -\partial^2}\right)^{-1/2} \left(\frac{\det \Delta_{\rm FP}}{\det -\partial^2}\right). \tag{16}$$

In the usual semi-classical method the dilute instanton gas [7] is noninteracting and one writes

$$\det D^2 = (\det d^2)^N, \tag{17}$$

where d^2 is the stability operator for the single monopole configuration.

We will see, however, that the d^2 has negative eigenvalues signifying the instability of a single monopole. In the background gauge (13) the operator d^2 has the form:

$$d^2 = \delta_{\mu\nu}\nabla_\alpha(\tilde{A}_{\rm cl})\nabla^\alpha(\tilde{A}_{\rm cl}) + i\,[\tilde{F}^{\rm cl}_{\mu\nu},\ ], \tag{18}$$

where $\tilde{A}_{\rm cl}$ is defined in (4) and $\tilde{F}^{\rm cl}_{\mu\nu}$ is the corresponding field. We refer to the literature for details of the stability analysis [9,1]. Except we mention that the source of the instability of d^2 is the long range Coulomb tail of the single regularized monopole solution.

It should also be noted that in the presence of Higgs fields the $1/r^2$ Coulomb tail is cancelled and we have a screened potential $e^{-m_H r}/r^2$, which removes the potential instability. m_H is the Higgs mass.

3.4 *Debye screening in the monopole gas*

Our main point is that when the instantons are interacting, the fluctuation problem cannot in general be factored into N copies of the fluctuation problem for an isolated monopole. Rather one should consider the stability of the neutral plasma of monopoles as a whole. This statement also applies to the Yang–Mills–Higgs system in the previous section. The main reason behind this is that we have an integration over the positions of the monopole and the charges, or equivalently a functional integration over the charge density field. The effect of this averaging over the charge density is very nontrivial. The results of [8] show that the charge density field clusters so that the theory of the density field generates a mass gap dynamically.

In fact the results of [8] mean that in the neutral plasma the fluctuations of the magnetic field are bounded and the $1/r^2$ magnetic field of a single monopole is Debye screened to e^{-Mr}/r^2, where $M(g^2, \lambda, z)$ is the non-perturbative mass gap, which depends on the coupling g^2, the monopole size λ which is effectively the cut-off of the Coulomb gas and the fugacity z. Recall that the source of instability for a single isolated monopole is the long range tail of the Coulomb potential. One might, therefore expect that in the screened neutral plasma a 'dressed' monopole whose Coulomb tail has been screened can in fact be stable.

In the following we shall assume that the monopole gas is dilute. Using translation invariance we focus on one monopole at $x = x_\alpha$ and its neighbourhood. We are thus considering the problem in the presence of a single source at $x = x_\alpha$. Recall that the fields outside the monopole cores of size λ are abelian. Thus, in the unitary gauge it follows from (7) and (4) that outside the core of this monopole the field is

$$\tilde{A}^{\rm out}_\mu = \sum_{a=1}^{N} \frac{1}{2} q_a \begin{pmatrix} D_\mu(x - x_a) & 0 \\ 0 & -D_\mu(x - x_a) \end{pmatrix}, \tag{19}$$

where we have set $x_N = x_\alpha$. Inside the core the effect of the core field of the other monopoles can be ignored and we have

$$\tilde{A}_\mu^{\text{in}} = \sum_{a=1}^{N-1} \frac{1}{2} q_a \begin{pmatrix} D_\mu(x-x_a) & 0 \\ 0 & -D_\mu(x-x_a) \end{pmatrix} + \frac{1}{2} \begin{pmatrix} qD_\mu(x-x_\alpha) & \tilde{W}_\mu^-(x-x_\alpha) \\ \tilde{W}_\mu^+(x-x_\alpha) & -qD_\mu(x-x_\alpha) \end{pmatrix}, \tag{20}$$

where $\tilde{W}_\mu^\pm \equiv \tilde{A}_\mu^1 \pm iq\tilde{A}_\mu^2$ and $\tilde{A}_\mu^1, \tilde{A}_\mu^2$ are as in (4). Introducing the charge density

$$\rho(x) \equiv \sum_{a=1}^{N} q_a \delta(x-x_a) \tag{21}$$

and assuming that in the mean, for large N, $\rho_N \simeq \rho_{N-1}$ we can rewrite (20) as

$$A_\mu(x;x_\alpha,[\rho]) = \int \mathrm{d}^3y \frac{\tau^3}{2} \rho(y) D_\mu(x-y) + \theta(\lambda - |x-x_\alpha|) \begin{pmatrix} D_\mu(x-x_\alpha) & W_\mu^-(x-x_\alpha) \\ W_\mu^+(x-x_\alpha) & -D_\mu(x-x_\alpha) \end{pmatrix}. \tag{22}$$

In the above expression the sharp θ function may be replaced by a smoother version.

Debye screening means that in the presence of a source the density $\rho(y)$ has a mean value $\bar{\rho}(y,x_\alpha)$, fluctuations around which are small. $A_\mu(x;x_\alpha,[\bar{\rho}])$ then represents a 'dressed' monopole configuration. In our case this 'source' is provided by the particular monopole at $x = x_\alpha$ in the plasma and the statement pertains to the field in the neighbourhood of this particular monopole.

The crucial point is that since Debye screening holds, we can assume self-consistently that the gas of 'dressed' monopoles is weakly interacting, unlike the 'bare' monopoles. The field around a dressed monopole decay exponentially over a distance scale set by the Debye screening length $l_D = 1/M$. If the average distance between the monopoles is much larger than l_D then the interaction between such dressed monopoles vanishes and the operator D^2 has a N-fold degeneracy. In other words the potential appearing in the stability equation resembles N far separated potential wells. In this situation we have, using translation invariance,

$$\det D^2 \simeq (\det D^2[\bar{\rho}])^N, \tag{23}$$

where $D^2[\bar{\rho}]$ now denotes the stability operator for a *single* dressed monopole which is the same for any monopole in the plasma. For finite distances between monopoles the exact degeneracy is lifted and eigenvalues of D^2 organize themselves in bands. This would lead to corrections to the result (23) which may be expanded in powers of l_D/l_m where l_m denotes the average distance between the monopoles.

In this regard there is a difference between the pure Yang Mills system and the Yang–Mills–Higgs system. The presence of the Higgs field means that the potential appearing in the stability operator around a single monopole decays over length scales of the order of $1/m_W$. Since the Debye length is much larger than $1/m_W$, corrections to the extensivity of the small fluctuation determinant appear as powers of $1/l_m m_W$.

3.5 *The Sine–Gordon transform and dynamical generation of mass gap*

We can now rewrite the theory in terms of a Sine–Gordon model. The path integral is written as

$$Z = \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{8g^6}{\sqrt{\pi}} \bar{s}^{3/2} \mathrm{e}^{-\bar{s}} \right)^N \sum_{\{q_a\}} \prod_{i=1}^{N} \int \mathrm{d}x_i$$

$$\exp\left(-\frac{2\pi}{g^2} \sum_{a \neq b} \frac{q_a q_b}{|x_a - x_b|} \right) (\Theta[\bar{\rho}])^N, \tag{24}$$

where we have defined $\Theta[\bar{\rho}]$ (functional of the charge densities) as

$$\Theta[\bar{\rho}] \equiv \left(\frac{\det D^2[\bar{\rho}]}{\det -\partial^2} \right)^{-1/2} \left(\frac{\det \Delta_{\mathrm{FP}}(\bar{\rho})}{\det -\partial^2} \right) \tag{25}$$

and $\bar{s}$ is the action for a single monopole.

It may be noted that we are dealing with a super renormalizable theory. Thus the expression (25) is ultraviolet finite.

We then have

$$Z = \sum_{N=0}^{\infty} \frac{1}{N!} \bar{J}^N \sum_{\{q_a\}} \prod_{i=1}^{N} \int \mathrm{d}x_i \exp\left(-\frac{2\pi}{g^2} \sum \frac{q_a q_b}{|x_a - x_b|} \right), \tag{26}$$

where the mean fugacity $\bar{J}$ is given by the formula

$$\bar{J} = \frac{8\, g^6}{\sqrt{\pi}} \bar{s}^{3/2} \mathrm{e}^{-\bar{s}} \Theta[\bar{\rho}]. \tag{27}$$

We can now perform the Sine–Gordon transform and we write (26) as

$$Z = \int \mathcal{D}\chi(x) \exp\left[-\frac{g^2}{32\pi^2} \int \mathrm{d}^3x \{ (\nabla\chi)^2, \quad -2\bar{M}^2(1 - \cos\chi) \} \right], \tag{28}$$

where

$$\bar{M}^2 = \frac{16\pi^2 \bar{J}}{g^2}. \tag{29}$$

It is important to emphasize in accordance with the discussion of [8], that the quadratic term $\int \mathrm{d}^3x (\nabla\chi)^2$ in (28) must be understood in a regularized sense, so that the Coulomb potential between the monopoles is valid only upto distances greater than the core size λ. In this sense λ is the cut off (lattice spacing) of the Sine–Gordon theory.

3.6 *Stability of the dressed monopole*

We now discuss, in some more detail than previously, whether the function $\Theta[\bar{\rho}]$ which is used in the definition of the fugacity $\bar{J}$ in (27) is well defined. This issue is important

because we have already indicated that as we average over $\rho(x)$ the operator $D^2[\rho]$ has negative eigenvalues when ρ corresponds to a single isolated monopole.

We will now argue that $D^2[\bar{\rho}]$ is a positive operator. Recall the form of $D^2[\bar{\rho}]$ in the unitary gauge, outside the core of the dressed monopole which we can choose to be at $x = 0$

$$D^2[\bar{\rho}] = -\delta_{\mu\nu}\nabla_\alpha(A(x,[\bar{\rho}]))\nabla^\alpha(A(x,[\bar{\rho}])) + i\,[\bar{F}^{(3)}_{\mu\nu}(x,[\bar{\rho}]),\ ], \tag{30}$$

where $\bar{F}^{(3)}_{\mu\nu}(x,[\bar{\rho}])$ is related to the magnetic field by $\bar{B}^{(3)}_\lambda = \epsilon_{\mu\nu\lambda}\bar{F}^{(3)}_{\mu\nu}$ and $\bar{B}^{(3)}_\lambda$ is given by

$$\bar{B}^{(3)}_\lambda = \frac{\partial}{\partial x_\lambda}\int \mathrm{d}^3y\,\frac{\bar{\rho}(y)}{|x-y|}. \tag{31}$$

Debye screening means that

$$\bar{B}^{(3)}_\lambda = \frac{x^\lambda}{r^3}\mathrm{e}^{-Mr}f(Mr), \tag{32}$$

where the function $f(x)$ has the property that for $x \gg 1$, $f(x) \sim 1$, and M is the mass gap related to the Debye length l_D by $M = 1/l_D$.

For a nonzero mass gap, the field outside the monopole core cannot be transformed to the radial form of (9). However close to the core and distances much smaller than the Debye length l_D the field is close to the single monopole field outside the core and may be cast in the radial gauge. Furthermore at distances much larger than the Debye length, the field is close to zero and once again one may cast the gauge potential in the radial gauge trivially (i.e. with $K(r) = 1$). As mentioned in § 4 the source of the even parity S-wave instability is the long range Coulomb field of the monopole. Since screening cuts off this Coulomb field and replaces it approximately by an exponential, one expects stability.

The situation is in fact similar to that of the Yang–Mills–Higgs system in some respects. Recall that the S-wave stability of the 't Hooft–Polyakov monopole is ensured by the fact the Higgs field rises *exponentially* to 1 beyond the core and cancels the negative $1/r^2$ tail of the gauge field, preventing the potential in the Schroedinger problem from being negative at large distances. In our problem the gauge field itself falls off exponentially to zero and thus the potential for the S-wave is positive at large distances. In this sense we have a dynamical Higgs effect.

3.7 *Self consistent dynamical generation of monopole size*

As mentioned before the fugacity of the monopole gas depends on $\bar{\rho}$ and hence it has a dependence on the mass gap M and the cutoff λ of the Sine–Gordon theory. This means that the cutoff λ is not independent, but determined in terms of M and g^2, i.e.

$$M = M(\lambda, g^2, z(\lambda, M)) \tag{33}$$

implies that

$$\lambda = \frac{1}{g^2}F\left(\frac{M}{g^2}\right), \tag{34}$$

where F is a function obtained by inverting (33). This equation now fixes the monopole core size λ as a function of the mass gap M in a self-consistent manner. The above considerations must be considered as qualitative because the calculation of the function F in (34) is beyond our present technology.

4. Dual representation and the disorder operator

We now relate the Euclidean formalism in terms of disorder operators introduced by 'tHooft [10] . In the hamiltonian formalism the Schroedinger picture disorder operator $\Phi_D(\mathbf{x}_0)$ is defined as an operator which creates a Z_2 magnetic vortex at the point $\mathbf{x}_0$ in two dimensional space. More specifically it implements a singular gauge transformation $\Omega^{[x_0]}$ which has the property that if we consider a closed spatial loop C parameterized by an angle θ one has

$$\Omega^{[x_0]}(\theta + 2\pi) = -\Omega^{[x_0]}(\theta) \tag{35}$$

when the loop C encloses the point $\mathbf{x}_0$. If C does not enclose $\mathbf{x}_0$ the gauge transformation is single valued. Consider now the two point function of the Heisenberg picture disorder operators

$$\langle \Phi_D^\dagger(x)\Phi_D(y)\rangle. \tag{36}$$

Here x and y stand for three dimensional coordinates (including the Euclidean time). This two point function is then a sum over all configurations of the gauge fields which have a Dirac string singularity along a line joining x and y with a monopole of charge $1/2$ at the point y and an anti-monopole of charge $-1/2$ at the point x. It is crucial that the magnetic charges of these monopoles is half that of the monopoles which populate the vacuum. They have magnetic charges so that the Dirac string is *visible* by the lowest electrically charged quarks which couple to the gauge field.

Repeating the steps which led to the Sine–Gordon representation with the difference that we have two external magnetic sources with charges $\pm 1/2$ we easily get

$$\langle \Phi_D^\dagger(x)\Phi_D(y)\rangle = \langle e^{i/2(\chi(y)-\chi(x))}\rangle, \tag{37}$$

the average on the right hand side in (37) being performed in the Sine–Gordon theory. A similar identification holds for all higher point correlation functions of the disorder operators. Hence we can identify the disorder operator with

$$\Phi_D(x) = e^{i/2\chi(x)} \tag{38}$$

In fact the Sine–Gordon action may be now written in terms of Φ_D as

$$S = \frac{g^2}{32\pi^2}\int d^3x[\partial_\mu \Phi_D^\dagger \partial_\mu \Phi_D + M^2((\Phi_D)^2 + (\Phi_D^\dagger)^2)] \tag{39}$$

up to an irrelevant constant. The field Φ_D is not a conventional scalar field, since $\Phi_D^\dagger \Phi_D = 1$. This non-linear Z_2 sigma model can be generalized to a linear sigma model by the addition of the term $\lambda \int d^3x(\Phi_D^\dagger \Phi_D - 1)^2$ to the action (39). The action then exactly has the form conjectured in [10].

The Sine–Gordon theory thus is itself a dual representation of the original Yang–Mills theory. The action (39) has the global Z_2 symmetry $\Phi_D \to \Phi_D^\dagger$ which is spontaneously broken leading to magnetic disorder and confinement. This is simply the symmetry $\chi \to \chi + 2\pi$ of the Sine–Gordon model. It is clear from the action (39) that the dimensionless coupling constant is $\sqrt{M}/g$. This is inversely related to the gauge coupling g as expected in a dual formulation. The dual theory is weakly coupled when M/g^2 is small. In this limit the minima of the potential $(\cos\chi)$ break the Z_2 symmetry spontaneously.

Finally we note that the above construction of the disorder operators can be easily extended to $SU(N)$ gauge theories following the treatment in [3].

5. Conclusions

We have argued that in $2+1$ dimensional pure Yang–Mills theory Debye screening in a gas of regularized and dressed Wu–Yang monopoles provides a consistent picture of quark confinement. We have used the results of [8] that in a three dimensional Coulomb gas the charge density field always clusters, leading to Debye screening even for arbitrarily low temperatures. Our line of argument has been self-consistent in nature, because Debye screening in turn implies a screened magnetic field and hence the stability operator around a single dressed monopole is expected to have no negative eigenvalues. The mass gap thus obtained is non-perturbative and determines the monopole size self-consistently. A related issue is that the mean configuration $\bar{\rho}(x)$ in the presence of a single monopole source is in general non-classical and hence the associated scalar potential $\bar{\chi}(x)$ does not satisfy a classical equation. Hence the explicit evaluation of the Wilson loops is not as easily done. However on general grounds the existence of a mass gap leads to qualitative conclusions that are similar to the case of the Yang–Mills–Higgs system. Finally we have obtained a representation of the disorder operators of the theory in terms of the Sine–Gordon field which leads to a dual representation of the gauge theory.

References

[1] S R Das and S R Wadia, *Phys. Rev.* **D53**, 5856 (1996)
[2] A M Polyakov, *Nucl. Phys.* **B120**, 429 (1977)
[3] S R Wadia and S R Das, *Phys. Lett.* **B106**, 386 (1981)
[4] G 't Hooft, *Nucl. Phys.* **B190**, [FS3], 455 (1981)
[5] T T Wu and C N Yang, in *Properties of matter under unusual conditions* edited by H Mark and S Fernback (New York, 1969)
[6] T Banks, R Myerson and J Kogut, *Nucl. Phys.* **B129**, 493 (1977)
[7] C Callan, R Dashen and D Gross, *Phys. Lett.* **B63**, 334 (1976)
[8] D C Brydges, *Comm. Math. Phys.* **58**, 313 (1978)
D C Brydges and P Federbush, *Comm. Math. Phys.* **73**, 197 (1980)
[9] T Yoneya, *Phys. Rev.* **D8**, 2567 (1977)
[10] G 't Hooft, *Nucl. Phys.* **B138**, 1 (1978); *Nucl. Phys.* **B153**, 141 (1979)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 81–92

CP symmetry and its violation in fundamental interactions

S D RINDANI
Physical Research Laboratory, Ahmedabad 380 009, India

Abstract. The concept of discrete symmetries in classical and quantum physics is reviewed. An account is given of *CP* violation observed in the K-meson system and of other experiments where *CP* symmetry has been tested. The present theoretical ideas on *CP* violation within the standard model, and problems needing extension of the model are described. Finally, ideas and experimental approaches to *CP* violation beyond the standard model are reviewed in brief.

Keywords. Discrete symmetries; CP violation; standard model.

PACS Nos 11.30; 12.10

1. The discrete symmetries *C*, *P* and *T*

Continuous symmetries, dependent on parameters which can be varied continuously, have their own significance in physics because they lead to conservation laws. Discrete symmetries, on the other hand, depend on a discrete set of parameters, and do not lead to conservation laws. They are nevertheless important in the study of the laws of physics. The simplest of discrete symmetries one can imagine are the space and time inversion symmetries, generally called parity (P) and time reversal (T). In the context of the special theory of relativity, these symmetries get linked with one more discrete symmetry, viz., particle-antiparticle interchange, usually called charge conjugation (C). C has basically nothing to do with space-time and is a kind of 'internal' symmetry. C, P and T thus play an important role in particle interactions at high energies, which are generally described using relativistic quantum field theory.

Of the symmetries, C, P and T, not all have the same significance for all fundamental interactions. Even though at one time C, P and T were all believed to be good symmetries, it has been known now for a long time that C and P are violated by weak interactions. Not only that, even the lower symmetry *CP* corresponding to a combination of C and P, has been found to be definitely violated in the K-meson system, but nowhere else. According to current ideas of cosmology, the universe started off with equal amounts of matter and anti-matter, and to understand the predominance of matter over anti-matter at the present time, a violation of *CP* at a fundamental level in the early universe is essential. However, C, P and T are all good symmetries of strong, electromagnetic and gravitational interactions as we know them today.

The combination of all three, viz., *CPT*, has a special significance. This combined symmetry is guaranteed to hold for all local interactions which also possess relativistic invariance by the so-called *CPT* theorem. This has a number of consequences for properties of antiparticles. For example, it implies the equality of masses and lifetimes of

particles and antiparticles, and requires their charges to be equal and opposite in sign. In addition, the CPT theorem implies that CP invariance and T invariance are equivalent, as also CP violation and T violation. Because of this interconnection, very often no distinction is made between CP and T. While CP makes sense only at high energies where antiparticles can be produced, it can be tested even at lower energies through tests of T invariance.

2. Discrete symmetries in classical and quantum mechanics

In classical mechanics, the definitions of physical quantities like energy, momentum, angular momentum, etc., decide their transformation properties under the space-time symmetries P and T. Thus momenta change sign under P and T, energy is invariant under P and T, angular momentum changes sign under T, but not under P. Depending on the nature of the forces, one can then determine whether or not the system is invariant under P and T.

For example, if the dynamics is determined by laws of electrodynamics, both the particle equations of motion as well as Maxwell's field equations are invariant under P and T, provided the fields are assigned appropriate transformation properties.

Normally, C does not enter the classical domain. However, one could define it as an operation which changes the charge of a particle, leaving other attributes the same. In that case classical electrodynamics is invariant under C, provided the fields change sign under C.

In quantum mechanics, symmetries are implemented by means of unitary (or anti-unitary) operators which commute with the Hamiltonian H, according to Wigner's theorem. P is represented by a unitary operator which satisfies $U_P^2 = 1$, and therefore has eigenvalues ± 1. Thus in a theory with parity invariance the eigenstates of H can be chosen to be a state with a definite parity, called even or odd according as the eigenvalue of P is $+1$ or -1.

The operator for T has to be anti-unitary rather than unitary as can be seen from the requirement that the time-evolution equation be form-invariant under T. The evolution equation

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{1}$$

when acted on by the time-reversal operator U_T gives

$$i\frac{\partial}{\partial(-t)}|\psi(-t)\rangle = H|\psi(-t)\rangle, \tag{2}$$

where

$$U_T|\psi(t)\rangle = |\psi(-t)\rangle, \tag{3}$$

provided U_T satisfies

$$[U_T, H] = 0 \tag{4}$$

and

$$U_T i = -iU_T. \tag{5}$$

Thus U_T has to be antilinear. This fact has many important consequences, as we shall see later.

While C has no place in non-relativistic quantum mechanics, it arises naturally in relativistic quantum mechanics, particularly, in the Dirac hole theory for antiparticles. In Dirac's interpretation, antiparticles are hole states in a sea of completely filled negative-energy states. This description does afford an operation of charge conjugation:

$$\psi(\mathbf{x},t) \rightarrow U_C\psi(\mathbf{x},t) = C\overline{\psi}^T(\mathbf{x},t), \quad (\overline{\psi} = \psi^\dagger\gamma^0), \tag{6}$$

where C is a matrix, and $\psi(\mathbf{x},t)$ is a column vector solving the Dirac equation. However, the real solution of the problem of negative-energy states involves the use of second quantization, and lies in the realm of quantum field theory.

In the quantum field-theoretic description of the discrete symmetries C, P and T, all fields undergo a unitary transformation.

$$\begin{aligned}
\psi(\mathbf{x},t) &\rightarrow U_C^{-1}\psi(\mathbf{x},t)U_C = \eta_C C\overline{\psi}^T(\mathbf{x},t) \\
\psi(\mathbf{x},t) &\rightarrow U_P^{-1}\psi(\mathbf{x},t)U_P = \eta_P P\psi(-\mathbf{x},t) \\
\psi(\mathbf{x},t) &\rightarrow U_T^{-1}\psi(\mathbf{x},t)U_C = \eta_T T\psi(\mathbf{x},-t).
\end{aligned} \tag{7}$$

Here C, P, T on the right-hand side are matrices, and η_C, η_P, η_T are phase factors, reflecting the inherent phase-arbitrariness in the definition of states in quantum mechanics. These phases may be chosen arbitrarily. The above illustration is for a spin-1/2 Dirac field, but other fields would have similar transformation properties. Invariance of a theory under the discrete symmetries means invariance of the action

$$S = \int \mathrm{d}^4x\mathcal{L}(\mathbf{x},t), \tag{8}$$

where $\mathcal{L}(\mathbf{x},t)$ is the Lagrangian density.

The phase factors appearing in the transformations can be determined from experiment, at least relative to one another, provided one assumes the corresponding invariance to hold. These phase factors are given the name 'intrinsic' C, P or T, as the case may be.

In simple theories, invariance under CP is simple to verify. However, in complicated interacting theories with several terms, the phases η_P, η_C have to be chosen carefully to ensure invariance of the action. Only in the case when no consistent choice of phases exists for which the action is invariant can one say that theory violates CP.

As for the combined transformation CPT, it is found that there is always a consistent choice of phases for CPT invariance so long as the theory is local and Lorentz invariant. This is guaranteed by the CPT theorem due to Pauli, Lüders and Schwinger [1].

3. CP violation in nature

To look at a bit of history, it was believed for a long time that nature respects all the discrete invariances C, P and T. So much so, that paradoxical situations arose because of this prejudice. An example is the famous τ-θ puzzle, which came about in the early 1950s, when the strange mesons τ and θ were found to have identical properties except

for the fact that they decayed into two different final states:

$$\tau \to \pi^+\pi^-\pi^0$$
$$\theta \to \pi^+\pi^-. \tag{9}$$

This meant that the intrinsic parity assignments of τ and θ had to be opposite, for the pions were known to have parity -1, and the decays were expected to proceed with zero orbital angular momentum. The other explanation, which was considered heretical was that τ and θ were the same particle K, but parity was violated in these weak decays, and therefore no definite intrinsic parity could be assigned to K. This suggestion of parity violation in weak interaction proposed by Lee and Yang [2] led to the experiment of Wu *et al* [3] which confirmed parity violation in the radioactive β decay of ^{60}Co.

The experiment consisted in looking at the distribution of decay electrons from a polarized ^{60}Co-source. A forward-backward asymmetry relative to the spin direction of ^{60}Co nuclei was an indication of parity violation, since the observable $\mathbf{s}_N \cdot \mathbf{p}_e$ changes sign under $P(\mathbf{s}_N \to \mathbf{s}_N,\ \mathbf{p}_e \to -\mathbf{p}_e)$, and should have zero average value if P is conserved.

Further evidence for P violation came from the experiment of Culligan *et al* [4] who showed the $e^+(e^-)$ had helicity $+\frac{1}{2}(-\frac{1}{2})$ in the decays

$$\mu^\pm \to e^\pm + \nu + \bar{\nu}. \tag{10}$$

Since the initial state was unpolarized, this was evidence for P violation, since helicity changes sign under parity.

Eventually, the $V-A$ theory of weak interactions was formulated by Marshak and Sudarshan [5], and by Feynman and Gell-Mann [6], which incorporated the maximal parity violation observed experimentally. This made use of the current×current form of the interaction:

$$\mathcal{L}_W = \frac{G_F}{\sqrt{2}} J_{w\mu} J_W^{\mu\dagger};$$
$$J_{W\mu} = \sum_i \bar{\nu}_i \gamma_\mu (1-\gamma_5) l_i + \sum_i \bar{u}_i \gamma_u (1-\gamma_5) d_i, \tag{11}$$

where the sum is over generations of quarks and leptons. (In writing this equation, mixing between generations has been neglected for simplicity). This form of the interaction also violated C, but in such a way that CP was left invariant.

For the experiments, this implied that if the ^{60}Co experiment could be repeated with anti-^{60}Co, the forward-backward asymmetry would have been exactly opposite in sign to that observed in ^{60}Co, so that

$$\langle \mathbf{s}_N \cdot \mathbf{p}_e \rangle + \langle \mathbf{s}_{\bar{N}} \cdot \mathbf{p}_{\bar{e}} \rangle = 0 \tag{12}$$

as required by CP invariance. Similarly, since the helicities of $e^\pm$ in $\mu^\pm$ decay are equal and opposite, CP is not violated.

The first evidence for CP violation was found in 1964 by Christenson *et al* [7]. They found that a definite state (the long-lived neutral K meson) can decay into two different CP eigenstates: the predominantly $CP = -1$ 3π state, and a very small (0.2%) admixture of the pure $CP = +1$ 2π state. Since the intrinsic CP assignments of the pions were fixed from strong interaction production processes, this meant that the weak-interaction decay

process does not conserve *CP* (direct *CP* violation) or that K_L is not an eigenstate of *CP* (indirect *CP* violation), or both. In fact, K_L, which is an eigenstate of the Hamiltonian, is a linear combination of the states K^0 and $\overline{K}^0$ produced by strong interaction:

$$|K_L\rangle = a|K_1\rangle + b|K_2\rangle; \tag{13}$$

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K}^0\rangle) \tag{14}$$

$$|K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle), \tag{15}$$

where K_1 and K_2 are *CP* eigenstates. Hence even in the latter case, the Hamiltonian is not invariant under *CP*.

Direct *CP* violation is accompanied by a change in strangeness by 1 unit ($\Delta S = 1$), whereas *CP* violation in K^0–$\overline{K}^0$ mixing corresponds to $\Delta S = 2$. Thus, it is possible that only indirect *CP* violation exists. Alternatively, if only a $\Delta S = 1$ *CP*-violating term appears in the Hamiltonian, a $\Delta S = 2$ *CP* violation can be induced in second order of perturbation theory. Theories permitting only $\Delta S = 2$ *CP* violation are called superweak theories. Experiments can in principle check for direct *CP* violation.

Present experiments have only seen definite evidence for indirect *CP* violation [8]:

$$\frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_L \to \text{all})} = (2.03 \pm 0.04) \cdot 10^{-3},$$

$$\frac{\Gamma(K_L \to \pi^0\pi^0)}{\Gamma(K_L \to \text{all})} = (9.14 \pm 0.34) \cdot 10^{-4}. \tag{16}$$

Usually what is presented is [8]

$$|\eta_{00}| = \left|\frac{A(K_L \to 2\pi^0)}{A(K_S \to 2\pi^0)}\right| = (2.259 \pm 0.023) \times 10^{-3},$$

$$|\eta_{+-}| = \left|\frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)}\right| = (2.269 \pm 0.023) \times 10^{-3}. \tag{17}$$

For direct *CP* violation to be present, $|\eta_{00}| \neq |\eta_{+-}|$ for which there is no evidence.

In addition to this *CP*-violating charge asymmetry has been observed in semileptonic K decay. The K-meson system is the only place where *CP* violation has been observed experimentally. However, *CP* violation has been looked for in other experiments as well, which have only provided upper limits. Other possible experiments in which *CP* violation has been tested are as follows.

3.1 *Electric dipole moments*

One of the important consequences of *CP* (or rather *T*) invariance is the vanishing of electric dipole moments of particles. For an elementary particle, the only vector to which a dipole moment, which is itself a vector, can be proportional to is its spin. Thus $\mathbf{d} = \alpha\mathbf{s}$. This interaction energy of the electric dipole moment with an electric field is

$$H_{\text{e.d.}} = -\mathbf{d} \cdot \mathbf{E} = -\alpha\mathbf{s} \cdot \mathbf{E}. \tag{18}$$

Since s changes sign under T, while E does not, this interaction violates T. In the relativistic form,

$$H_{\text{e.d.}} = i\alpha \int \mathrm{d}^3x \bar{\psi}\sigma_{\mu\nu}\gamma^5\psi F^{\mu\nu}, \quad (F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu) \tag{19}$$

which can be shown to violate CP, P and T. A generalization can be made to a 'weak' dipole interaction

$$H_{\text{w.d.}} = i\tilde{\alpha} \int \mathrm{d}^3x \bar{\psi}\sigma_{\mu\nu}\gamma^5\psi(\partial^\mu Z^\nu - \partial^\nu Z^\mu) \tag{20}$$

which also violates CP, in addition to P and T.

Electric dipole moments of various particles have been looked for without success, giving only upper bounds [8]. Weak dipole moments of leptons have also been tested for in experiments at the Z resonance at LEP and upper bounds have been obtained [9].

3.2 *Detailed balance in nuclear reactions*

T invariance predicts equal amplitudes for forward and backward reactions. This has been tested in several reactions, as for example,

$$^{24}\mathrm{Mg} + \alpha \rightleftarrows {}^{27}\mathrm{Al} + p. \tag{21}$$

So far no violation has been observed, and the experiments give upper bounds on effective T-violating parameters.

3.3 *Spin and momentum correlations*

Triple-products of spin and momenta, $\mathbf{p}_1 \times \mathbf{p}_2 \cdot \mathbf{p}_3$ or $\mathbf{s}_1 \times \mathbf{s}_2 \cdot \mathbf{p}_3$ etc. are odd under T. Hence naively, their observation would signal violation of T. However, this is true only if final-state interactions are sufficiently weak and can be neglected. This is because a T operation interchanges initial and final states, in addition to changing the sign of momenta and angular momenta. Forward and backward amplitudes are equal in magnitude only if final-state interactions (giving rise to so-called 'strong phases') can be neglected. This is the case, for example, with the decay $n \to pe^-\bar{\nu}$, where triple products have been looked for unsuccessfully.

Alternatively, in cases when the incoming state is a CP eigenstate, the observation of a momentum or spin correlation which is CP odd, rather than T odd, would unambigously signal CP violation.

4. Theoretical understanding of CP violation

Most current experimental results are explained quite remarkably well by the so-called standard model, developed by Glashow, Salam and Weinberg. This theory is based on local gauge invariance under the group $SU(2)_L \times U(1) \times SU(3)_C$. Under this symmetry the spin-1/2 quarks and leptons are arranged in left-handed doublets of weak isospin group $SU(2)_L$ and right-handed singlets. The quarks are triplets under $SU(3)_C$

representing 'colour' transformations. At present there are three known generations of quarks and leptons, which amounts to a replication of the left-handed doublets thrice. Thus, there are three 'up'-type quarks corresponding to a value $+\frac{1}{2}$ for the third component of weak isospin (called u, c and t quarks), and three 'down'-type quarks corresponding to $-\frac{1}{2}$ for the third component of weak isospin (called d, s and b quarks). The other basic ingredients are the 12 spin-1 gauge bosons γ, $W^{\pm}$ (with a mass of 80.2 GeV), Z^0 (with a mass of 91.1 GeV), and G_a ($a = 1, 2, \ldots, 8$), and a spin-0 massive Higgs boson which is not yet seen. The photon and the gluons are massless. The constituents of the standard model are displayed below:

$$\begin{aligned}
&\text{Quarks: } \begin{pmatrix} u_{i\alpha} \\ d_{i\alpha} \end{pmatrix}_L (u_{i\alpha})_R, \\
&\text{Leptons: } \begin{pmatrix} \nu_i \\ e_i \end{pmatrix}_L (e_i)_R, \\
&\text{Gauge bosons: } \gamma,\ W^{\pm}(80.2\,\text{GeV}),\ Z^0(91.1\,\text{GeV}), G_a, \\
&\text{Higgs boson: } H^0(> 58\,\text{GeV}).
\end{aligned} \tag{22}$$

Both the generation index i and the colour index α run from 1 to 3, and the colour index a for gluons runs from 1 to 8. Quarks, leptons and the massive gauge bosons and Higgs boson get their masses by the mechanism of spontaneous symmetry breaking. This also produces a mixing among different generations. As a result, the charged-current weak interactions, which are mediated by $W^{\pm}$, are described by a complex, unitary, 3×3 matrix, called the Cabibbo–Kobayashi–Maskawa (CKM) matrix. This matrix is totally arbitrary in the theory, and is only determined from experiment. It can be described by 4 parameters, 3 of which are like Euler angles $(\theta_1, \theta_2, \theta_3)$ describing rotations among the 3 generations, while the fourth one is a phase (δ) making some of the matrix elements complex. This complex phase is the cause of *CP* violation in the couplings of quarks to $W^{\pm}$. The smallest number of generations for which the phase can exist is three. (There would be no *CP* violation in a two-generation standard model).

The mixing gives rise to charged-current interactions of the quarks represented by the interaction Lagrangian

$$\mathcal{L}_{\text{c.c.}} = \frac{g}{2\sqrt{2}} \bar{u}_{i\alpha} \gamma^{\mu} (1 - \gamma^5) W_{\mu}^{+} d_{j\alpha} V_{ij} + \text{h.c.}, \tag{23}$$

where i, j represent the flavour indices, and V_{ij} is the ijth element of the CKM matrix.

How does this show up in the K-meson system? Indirect *CP* violation arises because of the phase in the amplitude for $K^0 \leftrightarrow \overline{K}^0$ transition. The quark structure of K^0 in $d\bar{s}$ and that of $\overline{K}^0$ is $\bar{d}s$. The amplitude for mixing of K^0 with $\overline{K}^0$ arises in the standard model from a fourth-order box diagram with exchange of a W^+W^- pair. The intermediate quark states involved are combinations of all the u quarks (up, charm and top), taken two at a time. The resulting amplitude is a sum of terms, the ij term being proportional to $V_{id}V_{is}^*V_{js}^*V_{jd}$. This complex quantity leads to a phase in the K^0–$\overline{K}^0$ mixing matrix, leading to indirect *CP* violation. For, then, one finds on diagonalizing the matrix, that K_L is not a *CP* eigenstate, but a linear combination of K_1 and K_2. Thus K_L contains an admixture of the *CP* even state K_1, permitting it to decay into 2π even if *CP* is conserved in the decay.

The magnitude of CP violation is determined by the quantity

$$J = \sin^2\theta_1 \sin\theta_2 \sin\theta_3 \cos\theta_1 \cos\theta_2 \cos\theta_3 \sin\delta \lesssim 3.5 \times 10^{-5} \sin\delta, \tag{24}$$

where $\theta_1, \theta_2, \theta_3$ are constrained by other experiments.

Direct CP violation also arises in the standard model through the so-called penguin diagrams, which are second-order in the weak interaction coupling. The amplitudes for K^0 and $\overline{K}^0$ decaying into 2π are relatively complex. Hence even if there is no CP violation in the mixing, so that $K_L = (|K^0\rangle - |\overline{K}^0\rangle)/\sqrt{2}$, K_L could still decay into 2π, because the matrix elements for $K^0 \to 2\pi$ and $\overline{K}^0 \to 2\pi$ are different.

Thus, the standard model can explain both direct and indirect CP violation. However, precise comparison with experiment is made difficult because the bound-state problem: The quark wave functions in the $K, \bar{K}$ are not known well enough.

The standard model predicts, similarly, CP violation in the B-meson system, i.e. mesons involving the b quark. This is expected to be somewhat larger, because the combination of mixing angles appearing in the B system are less constrained. This is seen from the so-called unitarity triangle of the CKM matrix. Unitarity gives the relations

$$\begin{aligned}
&\sum_{i=u,c,t} V_{id} V_{is}^* = 0 \quad \text{(strange } K \text{ mesons)},\\
&\sum_{i=u,c,t} V_{is} V_{ib}^* = 0 \quad \text{(strange } B_s \text{ mesons)},\\
&\sum_{i=u,c,t} V_{id} V_{ib}^* = 0 \quad \text{(nonstrange } B_d \text{ mesons)}.
\end{aligned} \tag{25}$$

Each relation represents a triangle in the complex plane, since the sum of 3 complex numbers vanish. The angles of the triangles are a measure of CP violation, the magnitudes of individual terms representing decay probabilities.

In the first two cases, one of the sides is very small compared to the other two, hence the triangles collapse to straight lines. In the third case all sides are comparable, though not large. Thus, CP violation is sizeable. However, B_d decay probabilities are low, and hence larger samples are needed. A substantial number of these will be produced at B factories, and it is hoped that CP violation would be observable.

The standard model predicts effects in all other systems which are too small to be detected in conceivable experiments. For example, quark dipole moments vanish at the two-loop level [10], and lepton dipole moments vanish at the three-loop level [11]. The estimates for dipole moments are:

$$\begin{aligned}
d_n &\approx 10^{-33} e\,\text{cm},\\
d_e &\approx 10^{-38} e\,\text{cm}.
\end{aligned} \tag{26}$$

Thus any other effects seen experimentally will be because of physics beyond the standard model.

If SM agrees well with experiments, is there need to go beyond SM? Quite apart from the usual reason given, that there are too many parameters in SM, there seem to be other reasons why SM is not adequate for CP violation. These are, the resolution of the strong

CP problem, and the understanding of the baryon-antibaryon asymmetry of the universe, or the so called problem of electroweak baryogenesis.

5. The strong CP problem

The vacuum of the strong interaction sector of the standard model, governed by the dynamics of the $SU(3)_C$ group, has a complicated structure: There are topologically distinct field configurations possible, all of which are degenerate. These distinct vacua $|n\rangle$ are characterized by a 'winding number' n, which takes all integral values. These vacua are not invariant under certain $SU(3)$ gauge transformations which change n. The correct gauge-invariant vacuum can only be a linear superposition of these vacua:

$$|\theta\rangle = \sum_n e^{-i\theta n}|n\rangle. \tag{27}$$

Each value of θ labels a different theory. For a given value of θ, if one calculates Green's functions, one finds that one has to sum over paths which connect the vacuum $|n\rangle$ to $|n+\nu\rangle$, for different values of ν. This sum amounts to adding a term

$$S_\theta = \theta\nu \equiv \theta \int d^4x \frac{g^2}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{a\mu\nu} \tag{28}$$

to the action. This term is odd under *CP* and *P*, and leads to 'strong' *CP* violation. It can therefore contribute to the neutron dipole moment, and can therefore be constrained by the strong experimental limit on the dipole moment. This requires θ to be smaller than about 10^{-9}. One can ask how such a small parameter arises. There is no answer, and this is one problem. But this is not all.

Suppose one assumes that for some reasons θ is small, there is still a problem – there is another contribution which changes θ. The axial vector $U(1)_A$ current

$$j^\mu_A = \sum_{i=1}^{N_f} \bar{q}\gamma^\mu\gamma_5 q \tag{29}$$

has an anomalous divergence

$$\partial_\mu j^\mu_A = 2N_f\left(\frac{g^2}{32\pi^2} F^a_{\mu\nu}\tilde{F}^{a\mu\nu}\right). \tag{30}$$

As a result one can show that under a chiral rotation of quarks through an angle α, the value of θ changes by $2N_f\alpha$.

Now, in defining the physical quark states in the theory, one has to make a change of basis which amounts to a chiral $U(1)$ rotation through an angle $\alpha = \arg\det M$, where M is the quark mass matrix, arising from spontaneous breaking of the electroweak gauge symmetry. Thus the observable value of θ is $\bar{\theta} = \theta + 2N_f(\arg\det M)$. This makes the problem worse, as now $\bar{\theta}$ has to be $\leq 10^{-9}$. Hence θ has to be tuned to lie close to $-2N_f(\arg\det M)$ to within 1 part in 10^9!

Most solutions proposed to the strong *CP* problem are beyond SM. The most popular one requires the introduction of spontaneously broken global $U(1)_{PQ}$ symmetry, as proposed by Peccei and Quinn in 1977 [12]. These theories have a light spin-0 particle

known as the axion, which has not been seen. There are theories which make the axion 'invisible' in laboratory experiments [13], but such theories are replete with cosmological and astrophysical implications which will really test the theory [14]. In particular, the axion is a candidate for dark matter.

6. Electroweak baryogenesis

It was pointed out by Sakharov in 1967 [15] that if the present baryon asymmetry of the universe were to be understood in terms of fundamental processes occurring in the early universe which started with zero net baryon number, then three requirements had to be met. Firstly, there had to be baryon number (B) violating processes. Secondly, these processes had to violate C and CP. Thirdly, the processes had to be out of equilibrium at some epoch. C violation is needed to produce a difference in baryon and anti-baryon production, whereas CP (or T) violation ensures that the forward and backward processes proceeded at different rates. If the B production processes were to be in equilibrium, any excess produced would eventually be washed out by the reverse process. Inequilibrium during the Hubble expansion guarantees that this probability of the reverse process is smaller.

Grand unified theories which unify electroweak and strong interactions at a high unification energy scale satisfy all these criteria at temperatures around 10^{15} GeV, and can in principle generate baryon excess [16]. However, it was realized that during the electroweak phase transition at $T \approx 10^2$ GeV, non-perturbative anomalous B-violating processes can wash out whatever excess was produced. These processes can also regenerate a baryon asymmetry [17]. However, indications are that the SM CP violation cannot give a sufficient baryon number [18]. This seems to call for a new mechanism of CP violation.

Alternative scenarios for baryogenesis, as for example, through lepton-number generation at an intermediate scale, have been proposed [19]. These nevertheless involve new sources of leptonic CP violation.

7. CP violation in extensions of SM

In general, extra particles can give rise to more complex couplings and sources of CP violation. This is true in models with extra scalars, extra fermions, or both. An example of such a model is the minimal supersymmetric standard model.

An interesting possibility is spontaneous CP violation. In this case CP is imposed on the Lagrangian, and CP violation arises because a scalar field has a complex vacuum expectation value on minimization of the potential. This requires a minimum of two doublets. However, to prevent unwanted Yukawa couplings a certain discrete symmetry is imposed. Then spontaneous symmetry breaking needs three doublets at the least.

Spontaneous CP violation is generally associated with the domain wall problem in cosmology. As the universe cools below the temperature corresponding to the scale of CP breaking, domains of different CP phases form. The surface energy density in the walls separating the domains is sizeable. Also, because domain walls are two-dimensional, the

rate of total energy decrease is only like the temperature T. Eventually the energy in the domain walls far exceeds the closure density ($\rho_0 \approx 10^{-46}\,\mathrm{GeV}^4$) of the universe:

$$\rho_{\mathrm{wall}} \sim \langle\phi\rangle^3 T \approx 10^{-7}\,\mathrm{GeV}^4. \tag{31}$$

A solution is that the *CP*-breaking transition occurs before an inflationary period. In that case, domain walls are irrelevant since we live in one domain which has got enormously stretched during inflation.

Another possibility often pursued is that *CP* is violated by 'soft' terms of mass dimension less than 4 in the real scalar field part of the Lagrangian. In that case the effect in other sectors of the theory are small and calculable.

8. A model-independent approach

Without reference to specific models, it is possible to study *CP*-violating experimental effects in terms of an effective Lagrangian [20]. In this relatively model-independent approach, one writes down all possible *CP*-violating couplings relevant to the experiment and consistent with Lorentz invariance. In case of fermion pair production in e^+e^- for example, these effective couplings will have electric and weak dipole moment terms. One could then parametrize possible experimental effects in terms of coefficients of these effective *CP*-violating couplings. Furthermore, the parameters could themselves be determined from various candidate theories. This would also enable confrontation of these theories with experiment.

Using an effective Lagrangian, one can make predictions for various experimental signatures of *CP* violation. Two special types of effects are event asymmetries and *CP*-odd correlations. An asymmetry of the type

$$\mathcal{A} = \frac{P(A \to B) - P(A^c \to B^c)}{P(A \to B) + P(A^c \to B^c)}, \tag{32}$$

where P denotes a transition probability, and A^c and B^c denote respectively states *CP* conjugate to A and B, would be a measure of *CP* violation. $\mathcal{A}$ has the advantage of being dimensionless and with absolute value less than or equal to 1. A *CP*-odd correlation is the expectation value of an observable which is odd under the *CP* transformation. While *CP*-odd correlations may not be as convenient as asymmetries described above, they may be easier to determine experimentally.

9. Conclusions

CP as an approximate symmetry of fundamental interactions is violated in weak interactions. The standard model can explain the observed *CP* violation in K mesons and make predictions for observable *CP* violation in the B-meson system. Any other *CP* violation would lie in the domain of extensions of the standard model. The current theoretical ideas on the solution of the strong *CP* problem and baryogenesis seem to need some extension of the standard model. Quite apart from this, it would be pragmatic to keep an open mind and use model-independent approaches to derive constraints on possible *CP*-violating couplings from experiment.

References

[1] W Pauli, in *Niels Bohr and the development of physics*, edited by W Pauli, L Rosenfeld and V Weisskopf (New York, McGraw-Hill, 1955)
G Lüders, *Dank. Mat. Fys. Medd.* **28**, 5 (1954)
J Schwinger, *Phys. Rev.* **82**, 914 (1951)
[2] T D Lee and C N Yang, *Phys. Rev.* **104**, 254 (1956)
[3] C S Wu, E Ambler, R W Hayward, D D Hoppes and R P Hudson, *Phys. Rev.* **105**, 1413 (1957)
[4] G Culligan, S G F Frank, J R Holt, *Proc. Phys. Soc. London* **73**, 159 (1959)
[5] R E Marshak and E C G Sudarshan, *Phys. Rev.* **D109**, 1860 (1958)
[6] R P Feynman and M Gell-Mann, *Phys. Rev.* **D103**, 193 (1958)
[7] J H Christenson, J W Cronin, V L Fitch and R Turlay, *Phys. Rev. Lett.* **13**, 138 (1964)
[8] Review of Particle Properties, *Phys. Rev.* **D50**, Part I (1994)
[9] OPAL Collaboration, P D Acton *et al, Phys. Lett.* **B281**, 405 (1992)
OPAL Collaboration, R Akers *et al*, *Z. Phys.* **C66**, 31 (1995)
ALEPH Collaboration, D Buskulic *et al, Phys. Lett.* **B297**, 459 (1992); **B346**, 371 (1995)
[10] E P Shabalin, *Sov. J. Nucl. Phys.* **36**, 575 (1982)
[11] M E Pospelov and I B Kriplovich, *Sov. J. Nucl. Phys.* **53**, 638 (1991)
M J Booth, Chicago preprint EFI-93-1 (1993)
[12] R D Peccei and H Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977); *Phys. Rev.* **D16**, 1791 (1977)
[13] J Kim, *Phys. Rev. Lett.* **43**, 103 (1979)
M A Shifman, A I Vainshtein and V I Zakharov, *Nucl. Phys.* **B166**, 493 (1980)
A P Zhitnitskii, *Sov. J. Nucl. Phys.* **31**, 260 (1980)
M Dine, W Fischler and M Srednicki, *Phys. Lett.* **B104**, 199 (1981)
[14] G G Raffelt, MPI Munich preprint, hep-ph/9502358, and references therein
[15] A D Sakharov, *Zh. Eksp. Teor. Fiz. Pis'ma* **5**, 32 (1967)
[16] M Yoshimura, *Phys. Rev. Lett.* **41**, 281 (1978)
S Dimopoulos and L Susskind, *Phys. Rev.* **D18**, 4500 (1978)
S Dimopoulos and L Susskind, *Phys. Lett.* **B81**, 416 (1979)
D Toussaint, S B Treiman, F Wilczek and A Zee, *Phys. Rev.* **D19**, 1036 (1979)
S Weinberg, *Phys. Rev. Lett.* **42**, 850 (1979)
[17] V A Kuzmin, V A Rubakov and M E Shaposhnikov, *Phys. Lett.* **B155**, 36 (1985)
[18] M E Shaposhnikov, *JETP Lett.* **44**, 465 (1986); *Nucl. Phys.* **B287**, 757 (1987); *Nucl. Phys.* **B299**, 797 (1988)
A I Bochkarev, S Yu Khlebnikov and M E Shaposhnikov, *Nucl. Phys.* **B329**, 493 (1990)
[19] M Fukugita and T Yanagida, *Phys. Lett.* **B174**, 45 (1986)
M A Luty, *Phys. Rev.* **D45**, 445 (1992)
P Langacker, R D Peccei and T Yanagida, *Mod. Phys. Lett.* **A1**, 45 (1986)
H Murayama, H Suzuki, T Yanagida and J Yokoyama, *Phys. Rev. Lett.* **70**, 1912 (1993)
A Acker, H Kikuchi, E Ma and U Sarkar, *Phys. Rev.* **D48**, 5006 (1993)
P J O'Donnell and U Sarkar, *Phys. Rev.* **D49**, 2118 (1994)
[20] For a review of this topic see S D Rindani, Proceedings of the Workshop on High Energy Particle Physics, Madras, 1994 edited by S Uma Sankar, *Pramana – J. Phys.* **45**, 263 (1995)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 93–107

Chaos and its implications

B BUTI
National Physical Laboratory, New Delhi 110012, India

Abstract. Starting with the very definition of chaos, we demonstrate that the study of chaos is not an abstract one but can lead to some useful practical applications. With the advent of some powerful mathematical techniques and with the availability of fast computers, it is now possible to study the fascinating phenomena of chaos – the subject which is truly interdisciplinary. The essential role played by fractals, strange attractors, Poincare maps, etc., in the study of chaotic dynamics, is briefly discussed. Phenomena of self-organization, coherence in chaos and control of chaos in plasmas is highlighted.

Keywords. Chaos; turbulence, Alfven waves; plasma.

PACS No. 52.35

1. Introduction

The topic of my lecture, namely Chaos, is truly interdisciplinary. It occurs in a variety of physical systems e.g., astrophysics, optics, condensed matter physics, plasmas, fluids etc. and also in biological systems, chemical systems, atmospheric sciences, biophysics, ecology, neural networks etc. The mathematical techniques to study chaotic phenomena in any of these systems are the techniques of nonlinear dynamics. However, I will confine myself to plasma systems. We plan to develop the subject by trying to answer the following questions: How exactly do we define a chaotic system? How and why do we study chaos? Are there any practical applications of chaotic phenomena? Chaotic dynamics vis-a-vis turbulence?

There is no unique definition of chaos. The simplest way it can be defined is as follows: If we map the trajectory of a particle in a given field and if the trajectory is not smooth but scattered, then the system can be chaotic. The scattered mapping, in fact, is a necessary condition for the system to be chaotic. Another necessary condition for a chaotic system is its extreme sensitivity to the initial conditions; in other words the two neighboring trajectories, with very slightly different initial conditions, should diverge exponentially. For example if $|X_a(0) - X_b(0)| = \epsilon$, then $|X_a(t) - X_b(t)| = \epsilon \exp(\alpha t)$ with $\alpha > 0$, is a must for a chaotic system (a and b are two neighboring trajectories). The latter can be more rigorously defined by means of Liapunov exponent which is defined as

$$\mathcal{L} = \lim_{\tau \to \infty} \frac{1}{\tau} \frac{D(\tau)}{D(0)},$$

where D is the physical distance between two neighboring trajectories in phase space. It turns out that $\mathcal{L} > 0$, is necessary as well as sufficient condition for the system to be

chaotic. Moreover $\mathcal{L}^{-1}$ gives the characteristic time for the onset of chaos. Since $\mathcal{L} > 0$ means an inability to specify the microscopic state of the system as it evolves, one immediately starts asking the following question: Is Liapunov exponent related to the entropy in any way? The answer is yes. In fact the KS entropy $= \Sigma$ all positive $\mathcal{L}$; Kolmogorov–Sinai (KS) entropy is a measure of loss of information of a dynamical system. A word of caution may be in order here. Chaos should not be equated simply with the disorder. It is more appropriate to consider chaos as a kind of order without any periodicity. In place of periodicity, chaotic system is characterized by scale-invariance property according to which the structure of the system does not change with the changes in scaling.

Historically Jules–Henry Poincare was the first person to recognize the chaotic behaviour of a dynamical system. He encountered this unexpected behaviour in his studies of three-body problems in planetary systems. The puzzling behaviour encountered by Poincare was simply because he was dealing with nonintegrable systems. We will elaborate on the nonintegrable systems, which lead to chaos, in the next section.

1.1 *Nonintegrable systems*

A system with N degrees of freedom, if it has N independent integrals of motion, is an integrable system. This is only a necessary condition for integrability. The necessary and sufficient condition for integrability is satisfied only if N independent constants

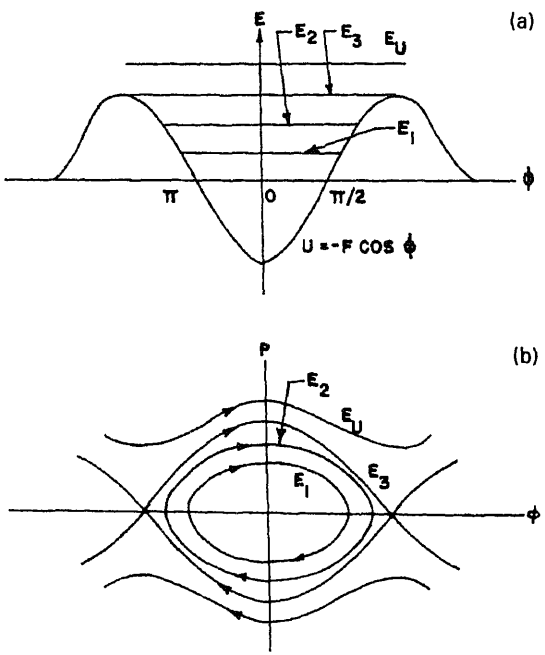


Figure 1. Energy diagram (**a**) and the corresponding phase space orbits (**b**) for an oscillator in a potential $U = -F \cos \phi$.

of motion $C_1, C_2, C_3, \ldots, C_N$ are in involution i.e., $\{C_i, C_j\} = 0$, where $\{\}$ represents the Poisson bracket. As a simple example, consider an oscillator moving in a potential $U = -F\cos\phi$ (cf. figure 1). The Hamiltonian for this oscillator, namely

$$H = p^2/2m + U = E,$$

has the total energy E which is conserved; this satisfies the necessary and sufficient condition for integrability. For $E < F$, the orbits are closed and for $E > F$, orbits are open but smooth. $E = E_s = F$, corresponds to the separatrix for which the period of oscillation is infinity. The two points, $\phi = \pm\pi$, where the two branches of the separatrix cross, are the hyperbolic points and the phase trajectories near these points diverge from them. Even very small perturbation in the neighborhood of the separatrix could lead an integrable system to nonintegrability. An integrable system can never lead to chaos whereas a nonintegrable system would.

1.2 *Surface of section*

Nonlinear dynamical systems are governed by nonlinear differential equations which most often can not be solved analytically. For such complex dynamical systems Poincare suggested geometrical method called surface of section. Let the phase space trajectories, of a particle or an oscillator in a given field, intersect the surface S at points $A_1, A_2, A_3, A_4, \ldots$. If the points $A_2, A_3, A_4, \ldots$ fall on A_1, then the system for sure is periodic. In case they do not fall on A_1 but the curve joining A_1, A_2, A_3 etc. is smooth, then the system is nonchaotic. On the other hand if the set of points A_1, A_2, A_3 etc. is random, the system is chaotic. We will use this technique for studying the chaotic phenomena in Hamiltonian as well as dissipative systems.

1.3 *Fractals and strange attractors*

In Hamilitonian systems, the phase space volume is conserved but in dissipative systems, the phase space shrinks continuously with increasing time; this leads to phase space of lower dimensionality and to an attractor. In case of regular motions the attractors are simple attractors e.g. a fixed point (sink) or a singly periodic orbit (limit cycle). For two-dimensional flows, sink and the limit cycle are the only two possibilities. However for three dimensional flows, one can also have doubly periodic orbits or attractors with complicated structures. Unlike the simple attractors like a point, a line etc. which have integral dimensions, attractors with complicated structures can have fractal dimensions; such attractors are known as strange attractors. Another characteristic, of strange attractors, is the scale invariance property i.e., on finer and finer scales the structure repeats itself. The attractor dimenson is defined as

$$D_a = \lim_{\epsilon \to 0} \frac{\ln n(\epsilon)}{\ln(1/\epsilon)},$$

where n is the minimum number of N-dimensional cubes of side ϵ needed to cover the basin of the attractor.

2. Nonlinear pendulum

As an illustrative example, let us consider the simplest possible nonlinear dynamical system namely a nonlinear pendulum, which is governed by the nonlinear equation

$$\frac{d^2\phi}{dt^2} + \omega_0^2 \sin\phi = 0, \tag{1}$$

with ϕ as the amplitude of the pendulum and ω_0 its angular frequency. For small ϕ, eq. (2) reduces to the equation of a simple pendulum whose solution is given by

$$\phi = a\cos(\omega_0 t) + b\sin(\omega_0 t). \tag{2}$$

Equation (2) is nonlinear but it can be solved exactly to give

$$\dot{\phi} = 2\mathcal{K}\, Cn(t;\mathcal{K}), \tag{3}$$

where Cn is the Jacobian elliptic cosine function and $\mathcal{K}^2 = (1+H)/2$ with H as the Hamiltonian corresponding to eq. (2) for $\omega_0 = 1$, i.e.

$$H = 1/2\,\dot{\phi}^2 - \cos\phi;\ \dot{\phi} = \frac{d\phi}{dt}. \tag{4}$$

The frequency of nonlinear oscillations is given by

$$(\pi/2)[1/F(\pi/2;\mathcal{K})], \tag{5}$$

where F is an elliptic integral of first kind. We may mention that the phase space diagram $(\phi, \dot{\phi})$ is identical to the one in figure 1 with $H = 1$ corresponding to the separatrix and $H < 1 (H > 1)$ leading to the trapped (untrapped) orbits.

Let us now consider a dampled nonlinear pendulum which is govened by

$$\frac{d^2\phi}{dt^2} + \sin\phi = -C\frac{d\phi}{dt}, \tag{6}$$

where C is the damping co-efficient. This damping could be due to air resistance provided by some sort of a vane attached to the swinging pendulum. The phase space diagram in this case shows a simple point attractor at $(0,0)$. If we drive this damped pendulum, say by a torque $(d\sin\phi)$, then this driven pendulum is described by the equation

$$\frac{d^2\phi}{dt^2} + C\frac{d\phi}{dt} + \sin\phi = -d\sin\phi\cos(\Omega t), \tag{7}$$

where Ω is the frequency of the driver. Numerical solution of eq. (7) shows the following interesting features [1]:

1. After the transit behaviour, the pendulum oscillates with a period which is double the period of the driver.
2. If we change the value of d slowly, for $C = 0.1$, we observe that at $d = 1.35$, there is a sudden jump in the pendulum frequency and momentarily it becomes singly periodic.
3. Further increase in $d(d > 1.35)$ leads to a sequence of period doubling bifurcations and eventually to chaos.

It is rather intriguing to find that a simple nonlinear dynamical system like a pendulum, when driven by an external torque, can become chaotic. Following the techniques used for the nonlinear pendulum, we can study the chaotic phenomena in all the other systems. We will further illustrate this by taking an example from plasmas. In this case we will also briefly discuss some applications of chaotic fields and say something about controlling the chaos.

3. Chaos in plasmas

The study of phenomena of chaos, in plasmas, is being pursued essentially to understand the phenomena of plasma turbulence which is very often observed in laboratory, space as well as in astrophysical plasmas. A chaotic system is essentially a turbulent system. As in the case of nonlinear pendulum, we have used the techniques of nonlinear dynamics to study the chaotic processes in plasmas. As an illustrative example, we will discuss chaotic Alfvén waves. Large amplitude Alfvén waves have been observed [2] in solar wind, planetary bow shocks, interplanetary shocks, environment of comets etc. Even the Alfvénic turbulence has been observed in the solar wind [3] as well as in the vicinity of comets [4].

A nonlinear wave, in general, is equivalent to a nonlinear dynamical system which exhibits the phenomena of chaos. Here we will discuss how and under what conditions, the nonlinear Alfvén waves can become chaotic; these chaotic fields in turn can lead to anomalous effects like plasma heating, particle acceleration and diffusion. These anomalous effects are essential to properly interpret some of the intriguing observations, e.g., solar coronal heating, energetic heavy ions in the vicinity of comets Halley and Giacobini-Zinner observed by recent cometary space missions [5, 6].

3.1 *Hamiltonion systems*

Let us consider a plasma which is embedded in a magnetic field $\mathbf{B}_0$ in the x–y plane. The electromagnetic waves propagating along x-direction in this plasma are governed by the two-fluid equations [7, 8]. By means of singular perturbation method, we can show that these fluid equations along with the generalized Ohm's law, lead to the following *evolution equation* for the nonlinear Alfvén waves [7, 8]:

$$\frac{\partial \mathbf{b}_\perp}{\partial t} + \alpha \frac{\partial}{\partial x}(\mathbf{b}_\perp \mid \mathbf{b}_\perp \mid^2) + \mu\left(\hat{e}_x \times \frac{\partial^2 \mathbf{b}_\perp}{\partial x^2}\right) = 0, \tag{8}$$

where $\mathbf{b}_\perp = (b_y, b_z)$, $\mu = V_A/(2\Omega_i)$, $V_A = B_0^2/4\pi\rho =$ Alfvén speed, $\Omega_i = eB_0/m_i c =$ ion cyclotron frequency, $\alpha = 1/4(1-\beta)$ with β as the plasma β i.e., the ratio of the kinetic pressure to the magnetic pressure. Equation (8) governs the elliptically polarized Alfvén waves; for circularly polarized waves this reduces to

$$\frac{\partial b_\pm}{\partial t} + \alpha \frac{\partial}{\partial x}(b_\pm \mid b_\pm \mid^2) \pm i\mu \frac{\partial^2 b_\pm}{\partial x^2} = 0, \tag{9}$$

where $b_\pm = b_y \pm i b_z$; b_+ for left hand polarized mode and b_- for right hand mode. Throughout this paper we have used the normalized units [7]. We may note that eq. (8) is

a coupled nonlinear equation whereas eq. (9) shows that left-hand and right-hand modes get decoupled. Equation (9) is the derivative nonlinear Schrödinger (DNLS) equation, which can be solved exactly; its solution is given by [9],

$$| b_+ |^2 = \pm \frac{8(1-\beta)(V-V_s)}{\sqrt{2}\cosh[2(V-V_s)(x-V_s t)] \pm 1}, \tag{10}$$

where V is the phase velocity of the linear Alfvén wave ($= V_A$ in unnormalized units) and V_s is the speed of solitary Alfvén wave. Equation (10) is the solitary wave solution for left-hand polarized Alfvén wave. Since the left-hand side of (10) is positive definite, the solution with (+) sign will correspond to

$$(1-\beta)(V-V_s) > 0$$

and the one with (−) sign to

$$(1-\beta)(V-V_s) < 0.$$

We may point out that the plasma β plays a very crucial role. For $\beta < 1$, super-Alfvénic solitons have larger amplitude compared to the amplitude of the sub-Alfvénic solitons, but for $\beta > 1$, the behaviour is reversed. It is interesting to note that for laboratory plasmas, $\beta < 1$ but for space plasmas, $\beta > 1$ for many systems.

So far, we have considered only uniform plasmas. However, very often, we encounter nonuniform plasmas. Repeating the procedure outlined above, we have derived the evolution equation for plasmas with weak but arbitrary inhomogeneity. The equation in this case is modified DNLS, which by means of a series of complicated transformations can be reduced to DNLS [10, 11]; once again we get solitary waves, which no more move with constant speed. These solitary waves get accelerated or decelerated depending on whether the wave propagation is along the increasing density gradients or decreasing density gradients. The evolution equation as well as its solutions are complicated, so they are not presented here (cf. ref. [10] and [11]).

Unlike (9), eq. (8) cannot be solved analytically. For any further analysis of eq. (8) and for investigating the possibility of chaotic behaviour of the system governed by this equation, it is useful to use the *Hamiltonian formulation.*

In a stationary frame of reference, namely,

$$\xi = (x - Vt)/\mu,$$

eq. (8) can be rewritten as

$$\hat{e}_x \times \frac{\mathrm{d}\mathbf{b}_\perp}{\mathrm{d}\xi} = \frac{\mathrm{d}\Psi}{\mathrm{d}\mathbf{b}_\perp}, \tag{11}$$

where

$$\Psi(\mathbf{b}) = \alpha\left[\frac{1}{4}(b^2 - b_0^2)^2 - \frac{\lambda b_0^2}{2}(\mathbf{b} - \mathbf{b}_0)^2\right]$$

with b_0 and λ as constants [7, 12, 13]. It is instructive to note that eq. (11) is equivalent to a classical equation of motion of a particle, with zero mass, which is moving under the influence of the coriolis force and a pseudo potential Ψ. The equipotential contours of Ψ, thus represent the trajectories of this pseudo particle with zero mass. In terms of the

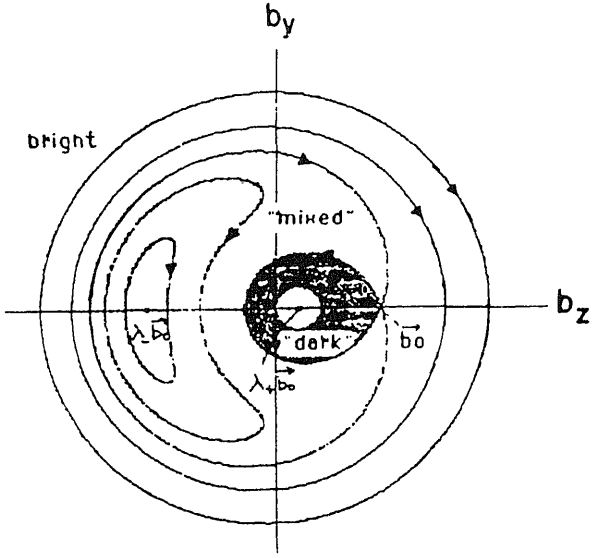


Figure 2. Equipotential contours of the pseudo potential Ψ. Also shown are 3 different solutions namely dark, bright and mixed solitons. Mixed solitons correspond to $\Psi < 0$.

Hamiltonian (H) which in this case is nothing but Ψ, eq. (11) can be represented by a set of following two equations:

$$\frac{db_y}{d\xi} = \frac{\partial H}{\partial b_z}, \tag{12a}$$

$$\frac{db_z}{d\xi} = -\frac{\partial H}{\partial b_y}. \tag{12b}$$

We immediately recognize these equations as Hamilton's equations with b_y and b_z as canonical co-ordinates. The potential Ψ being a quartic has four solutions, namely, the minimum at $(\lambda_- b_0)$, the maximum at $(\lambda_+ b_0)$ and the saddle point at b_0. For $0 < \lambda < 1/2$, $\lambda_\pm$ are given by [12],

$$\lambda_\pm \equiv \tfrac{1}{2}[-1 \pm (1 + 4\lambda)^{1/2}]. \tag{13}$$

For this range of λ, the equipotential contours of Ψ and the three solutions corresponding to dark, bright and mixed solitons are shown in figure 2. For $\beta < 1$, dark (bright) solitons have right (left) hand polarization. These solitons do not interact with each other and hence there is absolutely no possibility of having a chaos.

3.2 *Driven hamiltonian systems*

From figure 2, it is very apparent that the solitary (localized stationary nonlinear) Alfvén waves, left to themselves, will propagate as such for ever but this is somewhat an ideal situation. We would like to find out what happens to these solitary waves when there is

some local disturbance e.g., if there is a possibility of having another plasma wave generation or there is some solar activity (say a solar flare) which could propagate and interact with the Alfvén waves. To explore this, we will first consider the external driver as a plane wave. In order to make the problem tractable i.e., to be able to reduce the coupled partial differential equations (given by eq. (8)), modified by the driver, to a set of ordinary differential equations like eq. (12), we have to further restrict ourselves to the driver which is stationary in the frame of reference of the stationary Alfvén waves. This is indeed a very big restriction; in the latter part of this paper we will remove this restriction. Equation (12), in the presence of such a driver, is replaced by [7, 8],

$$\frac{db_y}{d\xi} = \frac{\partial H}{\partial b_z} + A\cos\theta, \tag{14a}$$

$$\frac{db_z}{d\xi} = -\frac{\partial H}{\partial b_y} + A\sin\theta, \tag{14b}$$

$$\frac{\partial\theta}{\partial\xi} = \Omega. \tag{14c}$$

In eqs (14), A is the amplitude of the driver and Ω is its frequency. We will now show that for $A \neq 0$, the dark, bright and the mixed solitons start interacting among themselves and when the driver is sufficiently strong i.e., A is sufficiently large, this interaction can lead to chaos. The reason for the appearance of chaos is rather transparent. In the absence of the driver, our system has two degrees of freedom (cf. eqs (12)) but in the presence of the driver, it has three degrees of freedom (cf. eqs (14)) and from our knowledge of nonlinear dynamics, we know that the minimum number of degrees of freedom required, for a system to be chaotic, is three and hence the possibility of chaos with the driver. For the left-hand polarized driver i.e., $\Omega = -2$, our numerical results are summarized by Poincare maps shown in figure 3. For $A = 0$, the entire set of Poincare points, originating from a given initial point, remains on the potential contour containing that initial point. For $A = 0.002$ i.e., for a weak driver, one of the sets of Poincare points, near the bright soliton separatrix, starts to scatter in a limited region of phase space. This leads to the onset of chaos for this set of Poincare points. We notice the formation of 3 islands near the local minimum of the potential Ψ. The number 3 is not a magic number; this is simply because we have taken $\omega/\Omega \sim 2/3$ (ω being the frequency of Alfvén waves). For somewhat stronger driver e.g., for $A = 0.2$, the system is almost chaotic except for the region inside the dark soliton (cf. figure 2). This region is unaffected because of the difference in polarity of the driver ($\Omega < 0$) and the Alfvén waves ($\omega > 0$). There is little interaction of the driver and the Alfvén waves because of opposite polarities. We repeated these calculations with the right hand polarized driver ($\Omega = 2$); these results are shown in figure 4. Comparison of Poincare maps for $A = 0.2$ for $\Omega < 0$ and $\Omega > 0$ testifies our above drawn conclusion. No matter how strong is the driver, the region inside the dark soliton remains coherent for $\Omega < 0$. It is this what we call 'coherence in chaos'.

3.3 *Dissipative systems*

Our entire discussion, so far, has been about the chaotic phenomena in nondissipative plasmas. Now we will very briefly present our efforts in this direction in connection with

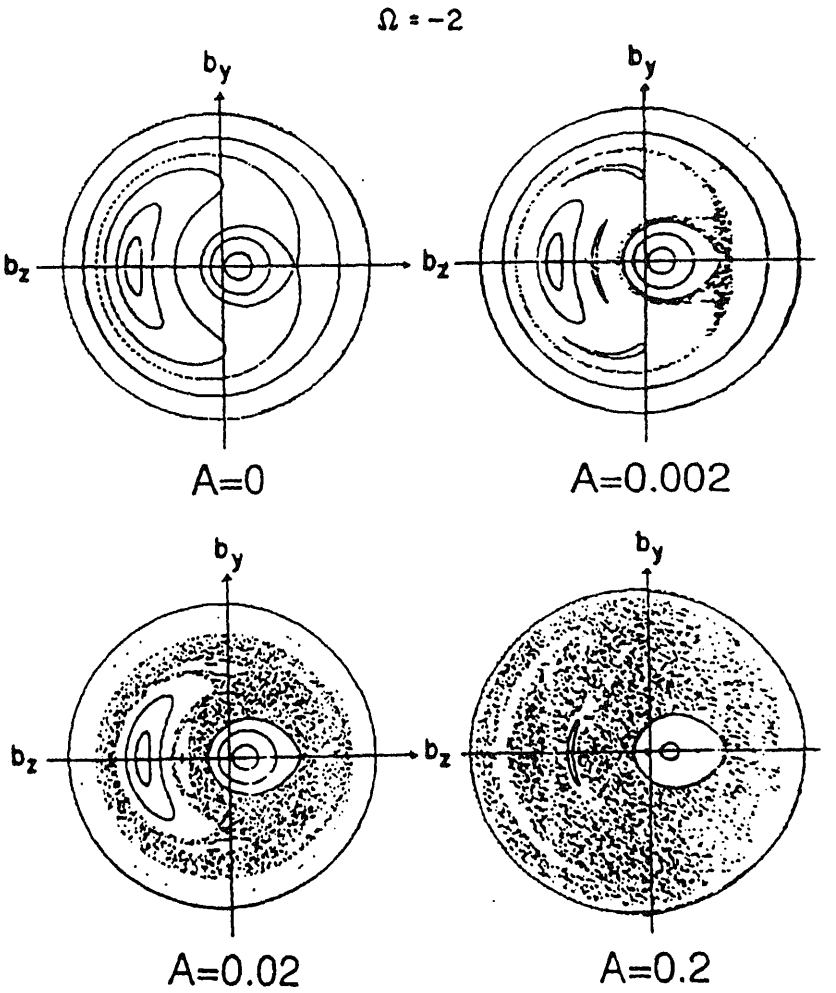


Figure 3. Poincare maps for the left hand polarized driver ($\Omega = -2$) for various amplitudes of the driver.

the dissipative systems. You notice that we had started with the simplest possible system and have been introducing the complexities into our system one by one. This step by step approach is simply to get a good physical insight into the complex phenomena of chaos.

For a dissipative two-species plasma, the nonlinear evolution equation for stationary Alfvén waves is given by [7, 13]

$$\hat{e}_x \frac{\mathrm{d}\mathbf{b}_\perp}{\mathrm{d}\xi} + \nu \frac{\mathrm{d}\mathbf{b}_\perp}{\mathrm{d}\xi} = \frac{\partial \Psi}{\partial \mathbf{b}_\perp} + A \begin{pmatrix} \cos \Omega\xi \\ \sin \Omega\xi \end{pmatrix}, \tag{15}$$

where $\nu > 0$ for dissipation and in the last term on right side of eq. (15), we have $\cos(\Omega\xi)$ for b_y and $\sin(\Omega\xi)$ for b_z component of the vector equation. Following the procedure outlined in the preceding sections, we have done the detailed analysis of eq. (15) and have

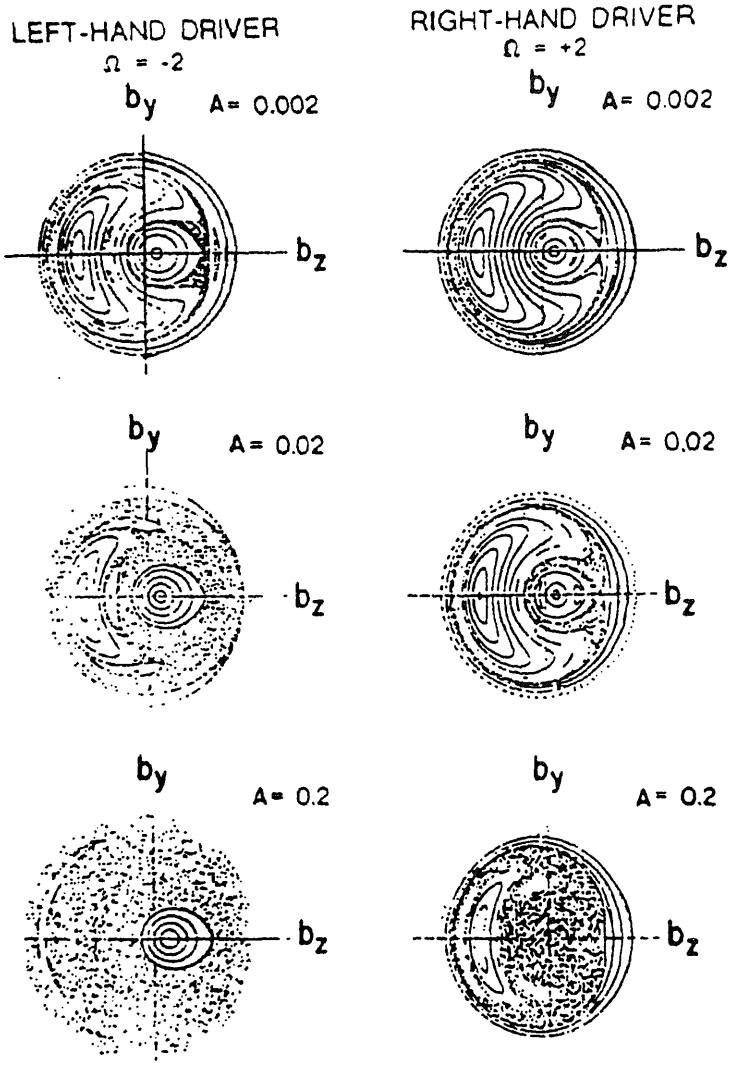


Figure 4. Comparison of Poincare maps for the left and the right hand drivers. The left hand column is the repetition of figure 3.

come up with some very interesting results. Summary of these results is presented in figures 5 and 6. Figure 5 shows the orbits as well as the Poincare points (shown by dots) in the $b_y - b_z$ phase space. The number of dots represents the periodicity. For $A = 0$, we found two attractors: one at the minimum of the potential and the other one at the saddle point. Figures 5 and 6 show the evolution of these attractors as A increases. The most interesting observation that we make is the complicated structure of the attractor for $A = 0.16$ (shown in I). In this box, we have not drawn the orbits but only the Poincare points are shown. This attractor is a strange attractor; this we have confirmed by checking its self-similar property and by determining its fractal dimension which

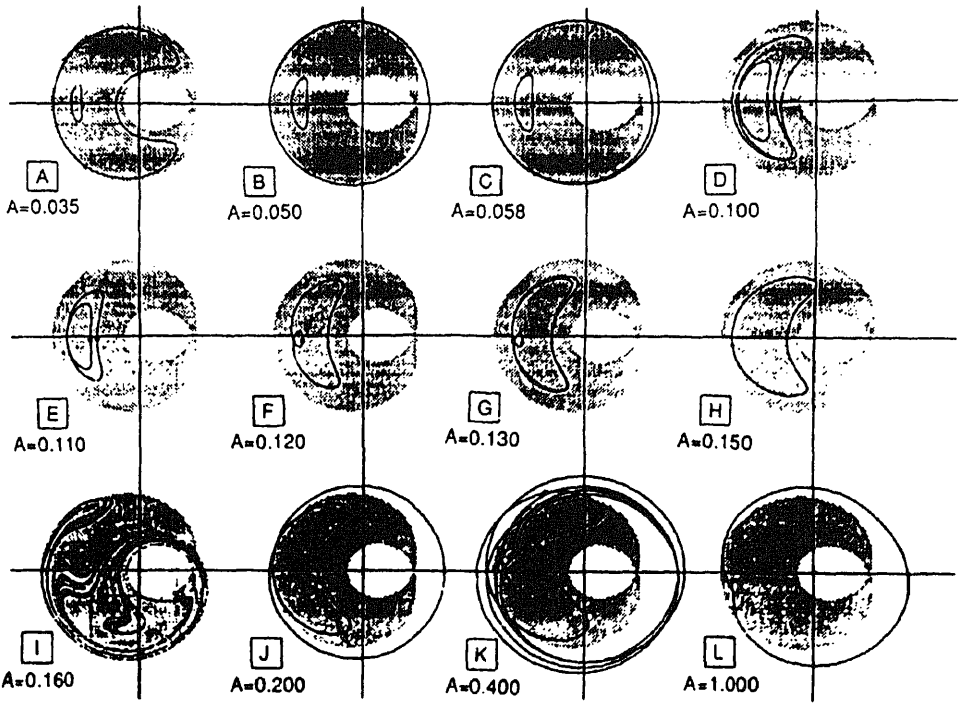


Figure 5. Shows attractors for the driven dissipative system in $b_y - b_z$ phase space. The labels are the same as in figure 6. The dots on the trajectories are the Poincare points. In (I), trajectories are omitted and only the Poincare points are plotted. The complicated structure of the attractor in this case represents a strange attractor for $A = 0.16$.

is found to be 1.57. For further discussion of figure 5, refer to Hada *et al* [13]. Another very interesting observation we make from figure 6 is that the chaos arises through two very distinct channels. At the saddle point, there are three sporadic attractors at $A \sim 0.035$, $A \sim 0.042$ and $A \sim 0.102$. The chaos in this channel arises through a sequence of period doubling bifurcations. However, at the minimum of the potential, in the region of strange attractor, there is a sudden transition to chaos. For more details, see Hada *et al* [13].

4. Control of chaos

So far our discussion has been confined to plasmas with only two species e.g., hydrogen plasma. Very often, we encounter multispecies plasmas with electrons, protons and heavy ions; the heavy ions, in laboratory plasmas, may be as impurities and in some natural plasmas as a genuine constituent, e.g., solar wind is composed of electrons, protons and α-particles (helium).

To study the chaotic processes in such multispecies plasmas, starting from the corresponding multifluid equations we derived the evolution equation [9] which is given

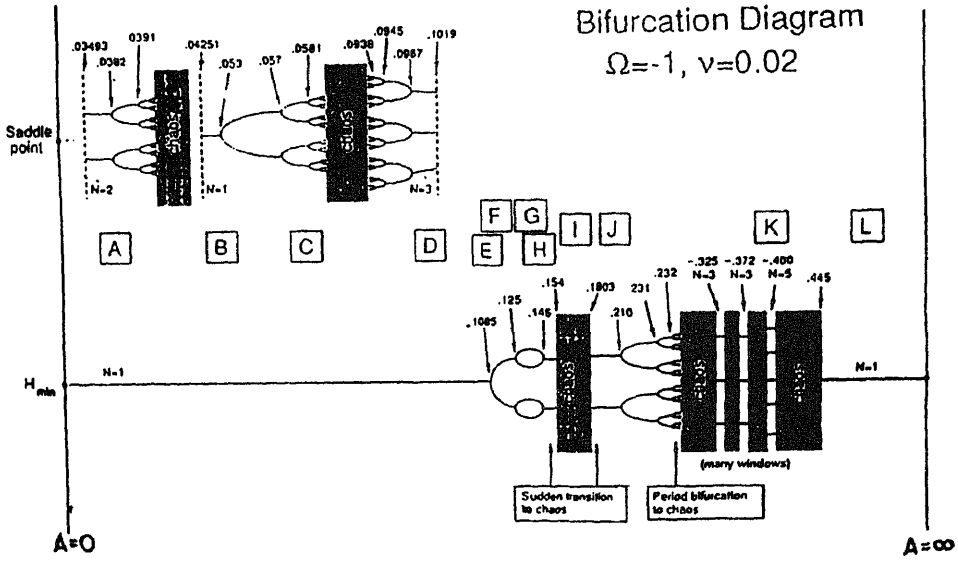


Figure 6. Shows the evolution of various attractors as the amplitude of the driver increases. The numbers on the top side correspond to the amplitudes. The top channel shows the normal as well as the reverse bifurcations. The bottom channel gives sudden transition to chaos in the neighbourhood of the strange attractor at $A = 0.16$.

by eq. (8) but with α and μ defined by

$$\alpha = \frac{1}{4}\sum_s \frac{\rho_{s0}^2(1 - Z_s\delta)}{(\rho_{s0} - \gamma_s P_s)},$$

$$\mu = \frac{1}{2}\sum_s \frac{\rho_{s0}}{Z_s},$$

$$\delta = \sum_s \frac{\rho_{s0} Z_s \gamma_s P_s}{(\rho_{s0} - \gamma_s P_s)} \left[\sum_s \frac{Z_s^2 \rho_{s0}^2}{(\rho_{s0} - \gamma_s P_s)}\right]^{-1}.$$

For solar wind parameters with 5% helium in abundance, we repeated our calculations; Poincare maps, for the two-species and the three-species plasmas, are shown in figures 7 and 8 respectively [14]. For both the figures, the driver is the left hand driver. Since all the parameters for both the figures are the same, figure 8 shows the effect of α particles. From comparison of these two figures, it is evident that the chaos is reduced due to the presence of helium–in other words, the threshold for chaos goes up because of heavy ions. Physically this could be interpreted as the inertial stabilization due to heavy ions.

5. Anomalous heating and particle acceleration

Chaotic systems can be very efficient sources of particle acceleration and plasma heating [15, 16, 17]. Lakhina and Buti have shown that the pick-up cometary ions [15] as well as the hot electrons injected from the solar coronal loops [17] can lead to unstable lower-

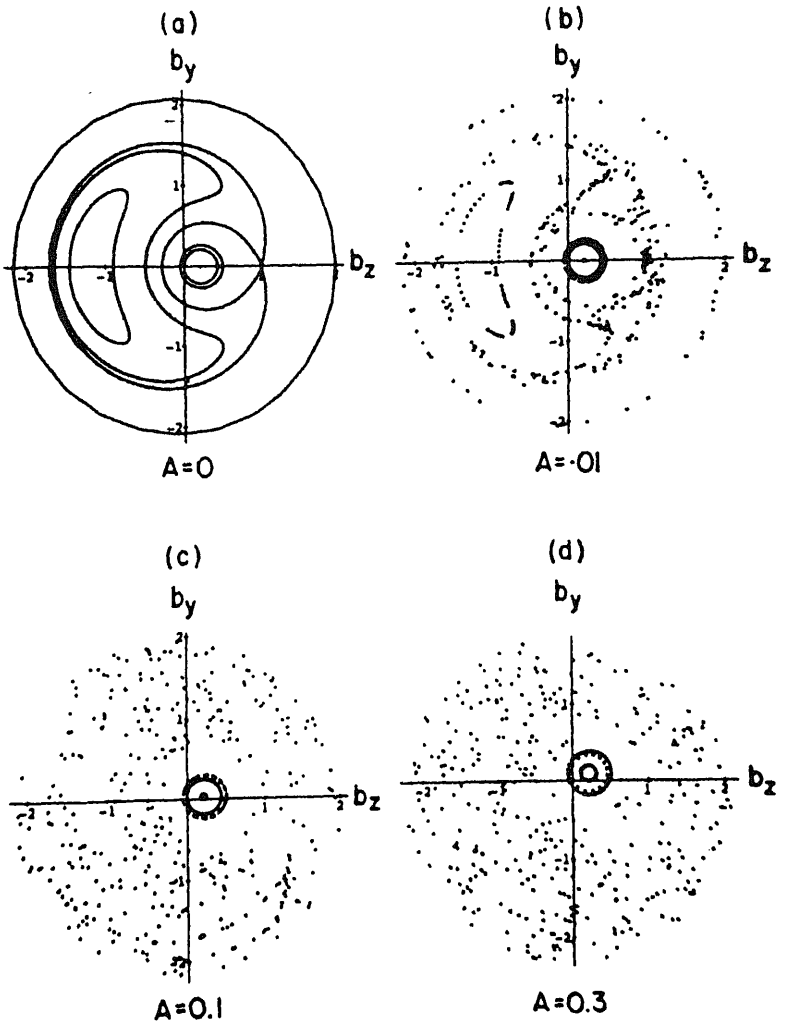


Figure 7. Poincare maps for driven Hamiltonian system for solar wind at 1 AU with only two species (electrons and protons). Driver in this case is left hand polarized.

hybrid waves. The saturated electric fields associated with these lower-hybrid waves, in both the cases, exceed the threshold for the generation of chaotic fields. They also showed that these chaotic fields preferentially accelerate the heavier ions. The maximum energy gained by the particles, from the chaotic fields, is proportional to the 5/3rd power of the mass of the particle. A part of the maximum energy gained is utilized in anomously heating the plasmas. We are now investigating the contributions of chaotic Alfvén waves, discussed in this paper, towards the anomalous heating of solar corona and anomalous acceleration of cometary ions. These will be reported in a forthcoming paper.

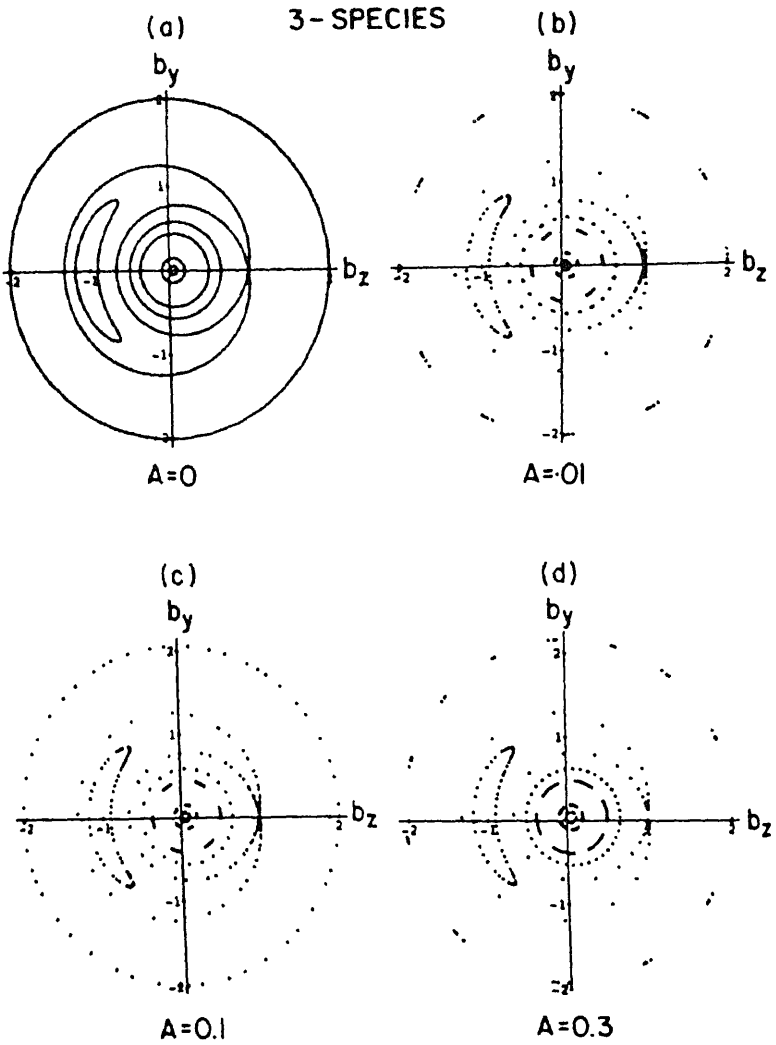


Figure 8. Same as figure 7 but for solar wind with three species i.e., electrons, protons and α particles; $N_\alpha/N_p = 0.05$. This figure is also valid for cometary plasma with water group ions as the third species and with $N_i/N_p = 0.01$ and $m_i/m_p = 16.8$ (i represents heavy ions and p represents protons).

6. Conclusions

The techniques, used in this paper for the study of chaotic Alfvén waves, can be used for the study of chaotic phenomena in any other system. Here we have discussed only the stationary waves with a very special kind of driver. However, this restriction on the driver can be relaxed; we can consider any other driver e.g., a beam representing solar wind or a pulse representing solar burst or a wave packet or any other type but then we cannot reduce our partial differential equation (cf. eq. (8)) to ordinary differential equations and

we have to solve partial differential equations. We now have a code for solving PDEs; this solution gives us the time series (time evolution) for b_y and b_z. From the time series, like in the case of observed time series, we have calculated the energy spectrum, correlation dimension and Lyapunov exponents which characterize the phenomena of chaos [18]. The latter calculations for a variety of drivers are under progress. The different spectra obtained for different drivers will be compared with the observed spectrum to finally conclude about the source of observed Alfvénic turbulence. As mentioned earlier, here we have considered Alfvén waves but the procedures outlined here are applicable to any other mode.

Acknowledgement

I would like to express my gratitude to all my collaborators of the contributions presented in this lecture.

References

[1] L Marvin and DeJong, *Phys. Teacher* **30**, 115 (1992)
[2] F L Scarf, F V Coroniti, C F Kennel, D A Gurnett and E J Smith, *Science* **232**, 382 (1986)
[3] L F Burlaga, *Rev. Geophys. Space Phys.* **21**, 363 (1983)
[4] B T Tsurutani and E J Smith, *Geophys. Res. Lett.* **13**, 259 (1986)
[5] S McKenna-Lawlor, E Kirsch, D O'Sullivan, A Thompson and K P Wentzel, *Nature* (*London*) **321**, 299 (1986)
[6] R J Hynds, S W H Cowley, T R Sanderson, K P Wentzel and J J Rooijen, *Science* **232**, 361 (1986)
[7] B Buti, *Solar and Planetary Plasma Physics* edited by B Buti (World Scientific, Singapore, 1990) 92
[8] C F Kennel, B Buti, T Hada and R Pellat, *Phys. Fluids* **31**, 1949 (1988)
[9] F Verheest and B Buti, *J. Plasma Phys.* **47**, 15 (1992)
[10] B Buti, *Geophys. Res. Lett.* **18**, 809 (1991)
[11] B Buti, *J. Plasma Phys.* **47**, 39 (1992)
[12] T Hada, C F Kennel and B Buti, *J. Geophys. Res.* **94**, 65 (1989)
[13] T Hada, C F Kennel, B Buti and E Mjolhus, *Phys. Fluids* **B2**, 2581 (1990)
[14] B Buti, *J. Geophys. Res.* **97**, 4229 (1992)
[15] B Buti and G S Lakhina, *Geophys. Res. Lett.* **14**, 107 (1987)
[16] B Buti, *Cometary and Solar Plasma Physics* edited by B Buti (World Scientific, Singapore, 1988) 221
[17] G S Lakhina and B Buti, *Solar Phys.* submitted (1995)
[18] L Nocera and B Buti, *Proceedings of ICPP 94* (1995)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 109–127

Nonlinear wave modulations in plasmas*

N N RAO
Physical Research Laboratory, Navrangpura, Ahmedabad 380009, India

Abstract. A review of the generic features as well as the exact analytical solutions of coupled scalar field equations governing nonlinear wave modulations in plasmas is presented. Coupled sets of equations like the Zakharov system, the Schrödinger–Boussinesq system and the Schrödinger–KDV system are considered. For stationary solutions, the latter two systems yield a generic system of a pair of coupled, ordinary differential equations with many free parameters. Different classes of exact analytical solutions of the generic system which are valid in different regions of the parameter space are obtained. The generic system is shown to generalize the Hénon–Heiles equations in the field of nonlinear dynamics to include a case when the kinetic energy in the corresponding Hamiltonian is not positive definite. The relevance of the generic system to other equations like the self-dual Yang–Mills equations, the complex KDV equation and the complexified classical dynamical equations is also pointed out.

Keywords. Nonlinear waves; modulational instability; solitons; NLS equation; KDV equation; Hénon–Heiles Hamiltonian; integrability; complexification.

PACS No. 52.35

1. Introduction

Coupled second-order ordinary differential equations for a pair of real scalar fields occur in many branches of physics. For example, in particle physics, the relativistic quantum field theories in $1+1$ dimensions for finite-energy localized fields lead to a set of nonlinear equations for a pair of scalar fields [1]. On the other hand, in plasma physics, the nonlinear development of the instability associated with the envelope modulation of an high-frequency wave packet coupled to an appropriate low frequency wave field is governed [2] by a pair of coupled equations like Zakharov equations, or the Schrödinger–Korteweg-de Vries (KDV) (or, –Boussinesq) system. For stationary solutions, the latter two systems lead to a generic system of a pair of coupled ordinary differential equations. Depending on the problem at hand, the generic system has varying number of free parameters which can take values over a wide range. An outstanding mathematical problem associated with such nonlinear coupled equations is to obtain their exact analytical solutions valid over as much of the parameter space as possible.

The generic equations are of interest in other contexts also. Recently, it was pointed out [3] that the stationary set derived from the Schrödinger–KDV (or, –Boussinesq) system is not only very similar to but also generalizes the well-known Henon–Heiles system [4]

* An extended version of this article appeared in *Pramana – J. Phys.*, Vol. 46, No. 3, 161–202 (1996)

which has been extensively studied over the last decade in the field of nonlinea dynamics. On the other hand, there is the interesting possibility that the self-dual Yang Mills field equations [5] could be related to the stationary equations derived from th Schrödinger–KDV (or, –Boussinesq) system. Furthermore, when the dependent variabl of an uncoupled KDV equation is made complex [6, 7], the resulting set of real equation are structurally very similar to those contained in the generic system. The latter reflect also certain interesting properties of the coupled equations obtained by the 'complex ification' of the dynamical equation for a classical particle with one-degree of freedom i a general conservative potential field [6].

In this review, I will briefly discuss some of the above features. In the process, I wi describe a method of solution that can be used for obtaining analytical solutions c coupled scalar field equations. In particular, I will use this method to obtain differe classes of exact analytical solutions of the generic equations obtained from th Schrödinger–KDV (or, –Boussinesq) system. Wherever possible, the examples are chose from the field of plasma physics where such equations are commonly encountere Unless otherwise stated, all the variables are assumed to be suitably normalized so th the equations are dimensionless throughout.

2. Modulational instability

Consider the propagation along $\hat{x}$-direction of a plane wave of frequency ω_0 a wavenumber k_0 represented by

$$\mathcal{E} = E(x,t)\exp[i(k_0 x - \omega_0 t)], \qquad (1$$

where the complex amplitude $E(x,t)$ is a slowly varying function of the space and t time variables. In order to derive a governing equation for the wave propagation, we st with the nonlinear dispersion relation, namely, $\omega = \omega(k, |E|^2)$ and Taylor expand around (ω_0, k_0) to obtain

$$\omega - \omega_0 = (k-k_0)\frac{\partial\omega}{\partial k} + \frac{1}{2}(k-k_0)^2\frac{\partial^2\omega}{\partial k^2} + |E|^2\frac{\partial\omega}{\partial(|E|^2)}, \qquad (2$$

where the partial derivatives have to be evaluated at $k = k_0$ and $|E| = 0$. Replacing t frequency shift $(\omega - \omega_0)$ by $i\partial/\partial t$ and the wavenumber shift $(k - k_0)$ by $-i\partial/\partial x$, obtain the following evolution equation for the slowly varying complex amplitude $E(x$

$$i\left(\frac{\partial E}{\partial t} + V_g\frac{\partial E}{\partial x}\right) + P\frac{\partial^2 E}{\partial x^2} + (Q\,EE^*)E = 0, \qquad (3$$

where asterisk denotes complex conjugate, $V_g \equiv \partial\omega/\partial k$ denotes the group veloci $P \equiv (\partial^2\omega/\partial k^2)/2$ represents the group dispersion, and $Q \equiv -\partial\omega/\partial(|E|^2)$ is the nonlin coefficient. In view of its structural similarity, eq. (3) is called the 'nonlinear Schrödin equation', and is known to govern the evolution of different types of high-frequer waves in plasmas. Examples are the Langmuir waves, the upper-hybrid waves and electromagnetic waves.

Equation (3) can be analysed to determine the stability properties of an high-frequer carrier wave when its amplitude is modulated with a lower frequency. Depending on

relative sign between the dispersive and the nonlinear terms, the modulation can become unstable. Linear stability analysis shows [2] that for $PQ > 0$ the waves are modulationally unstable. Physically, the instability arises due to the self-trapping of the wave field in the 'potential', which is determined by the wave itself [cf. the last term on the left-hand side of eq. (3)].

We shall now look for solutions of eq. (3) which are stationary in a frame moving with a constant speed. For this, we note that the second term in eq. (3) can be eliminated by going into a Galilean frame defined by $\zeta = x - V_g t$ and $\tau = t$:

$$i\frac{\partial E}{\partial \tau} + P\frac{\partial^2 E}{\partial \zeta^2} + (Q\,EE^*)E = 0. \tag{4}$$

We now look for stationary solutions in the (ζ, τ) coordinates by defining the variable, $\eta = \zeta - M\tau \equiv x - (V_g + M)t$ where M is the normalized speed, called the 'Mach number', of the stationary frame. In order to allow for any possible shift in the frequency as well as in the wave number of the carrier wave, we represent the amplitude field as

$$E = E_a(\eta)\exp[i\{X(\zeta) + T(\tau)\}], \tag{5}$$

where $E_a(\eta)$ is the real, stationary amplitude of the modulated wave. Substituting the solution (5) into eq. (4), we obtain from the imaginary part, $X(\zeta) = M\zeta/2P$ whereas the real part yields the following equation for the amplitude E_a:

$$2P\frac{\mathrm{d}^2 E_a}{\mathrm{d}\eta^2} = \lambda E_a - 2QE_a^3, \tag{6}$$

where $\lambda = 2(\mathrm{d}T/\mathrm{d}\tau) + (M^2/2P)$ is the nonlinear shift parameter. For localized boundary conditions, eq. (6) can be easily integrated to obtain the so-called 'envelope soliton', solution, namely,

$$E_a(\zeta, \tau) = \pm\left(\frac{\lambda}{Q}\right)^{1/2} \mathrm{sech}\left\{\left(\frac{\lambda}{2P}\right)^{1/2}(\zeta - M\tau)\right\}. \tag{7}$$

Clearly, the total high-frequency field $\mathcal{E}(x, t)$ has a structure wherein the amplitude of the carrier wave field is modulated, leading to a bell-shaped profile and the structure itself propagates with a constant velocity with respect to the laboratory frame.

Before considering further generalizations of the above equation, let us briefly summarize the physical mechanism that leads to such solutions. In plasma physics, the variable E can be taken to represent the electric field of a suitable high-frequency field. It can be easily shown that when the amplitude of such a wave field is slowly modulated spatially, the motion of a particle of charge q and mass m can be decomposed into two parts: the first part contains the linear motion that oscillates with the same frequency as the wave field, and is charge and mass dependent. The second part is the nonlinear motion arising because of the time-averaged force, and is independent of the sign of the charge. Such a nonlinear force is called the 'ponderomotive force' and is, in general, given by [8]

$$\vec{F}_p = -\frac{q^2}{4m\omega^2}\nabla(E_a^2).$$

Because of the inverse dependence on the mass, the ponderomotive force acts strongly c the electrons pushing them away from the regions where the field is stronger. Howeve the resulting ambipolar field causes the ions to follow the electrons, thereby creating density well which further traps the high-frequency field. This process is continued till dynamic balance between the nonlinear and the dispersive effects is achieved. Th envelope soliton solution is simply a representation of such a state.

3. Coupled equations

The ponderomotive force due to a high-frequency field in a plasma drives low-frequenc oscillation which may be a normal mode of the system. For example, the amplitud modulated Langmuir oscillations are coupled to the wave-excited, low-frequenc acoustic fluctuations called the 'ion-acoustic waves'. The nonlinear Schrödinger equatic derived above considers, however, only the static response of the latter waves. Such a approximation can be justified when the envelope packet is nearly static. However, whe the envelope moves with finite, non-zero speed, dynamic response of the low-frequenc wave should be taken into account.

3.1 *Zakharov system*

For the Langmuir-ion-acoustic waves, Zakharov [9] suggested the following pair coupled equations:

$$i\epsilon\frac{\partial E}{\partial t}+\frac{3}{2}\frac{\partial^2 E}{\partial x^2}=\frac{1}{2}(\delta n_e)E, \tag{}$$

$$\frac{\partial^2}{\partial t^2}(\delta n)-\frac{\partial^2}{\partial x^2}(\delta n)=\frac{\partial^2}{\partial x^2}\left(\frac{1}{4}EE^*\right), \tag{}$$

where ϵ is a known small, real parameter, and δn is the perturbed ion number densi Under the approximation used to derive the above equations, δn is equal to the electr number density perturbation, namely, δn_e. Equation (8) is a Schrödinger-like equati with δn_e as the potential, which in turn is governed (since $\delta n_e = \delta n$) by the linear wa equation (9) driven by a source term involving the 'wave function'. The coupled set equations (8) and (9) is called the 'Zakharov system', and has been very extensive studied in the literature. It may be noted that the nonlinear Schrödinger equation can recovered from the Zakharov system if, under the assumption of static response, o neglects the time derivative term in the driven wave equation (9). However, for dynam response, we obtain, using the stationary variable $\xi \equiv x - Mt$,

$$\delta n=-\frac{1}{4}\frac{EE^*}{(1-M^2)}, \quad \phi=\delta n+\frac{|E|^2}{4}\equiv M^2\delta n, \tag{1}$$

where ϕ is the self-consistent ambi–polar potential. Thus, the solutions of the nonlin Schrödinger equation are valid when the Mach number M tends to zero.

Equations (8) and (9) (together with the quasi–neutrality condition, $\delta n_e = \delta n$) can integrated for stationary solutions to yield

$$E_a(\xi) = \pm [8\lambda(1 - M^2)]^{1/2} \mathrm{sech}(\mu\xi), \tag{11}$$

$$\phi(\xi) = -2\lambda M^2 \mathrm{sech}^2(\mu\xi), \tag{12}$$

where $\mu^2 = \lambda/3$, and the nonlinear shift parameter given by $\lambda = 2\Delta + (\epsilon^2 M^2/3) \approx 2\Delta$ is required to be positive so that the solutions (11) and (12) satisfy the localized boundary conditions. Equation (11) then requires $M^2 < 1$ for real E_a; that is, the high-frequency envelope wave packet 'loaded' with the low-frequency density fluctuations can only travel at sub-sonic ($M < 1$) speeds. Both the solutions are symmetric with respect to $\xi = 0$. Note that the solutions (11) and (12) imply the following relation:

$$\frac{E_a^2}{4} = -\frac{1 - M^2}{M^2}\phi. \tag{13}$$

Clearly, for a given envelope field amplitude E_a, the variable ϕ becomes large when $M^2 \to 1$, in which case the driven linear wave equation (9) needs to be replaced by a suitable nonlinear equation. This is discussed in the next section.

3.2 *Schrödinger–Boussinesq system*

The nonlinear Schrödinger equation as well as the Zakharov equations take into account only the linear response in the low-frequency dynamics. However, as pointed out above, for the so-called 'near-sonic' ($M \sim 1$) propagation, the amplitude of the low-frequency perturbation can be quite large requiring a nonlinear dynamical equation in place of eq. (9). For the coupled Langmuir and ion-acoustic waves, Makhankov [10] suggested the following coupled Schrödinger–Boussinesq equations as the appropriate set:

$$i\epsilon \frac{\partial E}{\partial t} + \frac{3}{2}\frac{\partial^2 E}{\partial x^2} = \frac{1}{2}(\phi)E, \tag{14}$$

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\phi - \frac{\partial^4 \phi}{\partial x^4} - \frac{\partial^2}{\partial x^2}(\phi^2) = \frac{\partial^2}{\partial x^2}\left(\frac{EE^*}{4}\right). \tag{15}$$

The latter equation is called the (driven) 'Boussinesq equation', and generalizes the linear wave equation (9) to include the dispersive effects (third term on the left-hand side) as well as the nonlinear effects (the fourth term).

For stationary solutions of the form

$$E = E_a(\xi)\exp[i\{X(x) + T(t)\}], \quad \phi = \phi(\xi),$$

where $\xi = x - Mt$, eqs (14) and (15) yield

$$3\frac{\mathrm{d}^2 E_a}{\mathrm{d}\xi^2} = \lambda E_a + \phi E_a, \tag{16}$$

$$\frac{\mathrm{d}^2 \phi}{\mathrm{d}\xi^2} = -(1 - M^2)\phi - \phi^2 - \frac{E_a^2}{4}. \tag{17}$$

Equations (16) and (17) have two free parameters λ and M defined above. They describe the stationary, bi-directional propagation of coupled Langmuir and ion-acoustic waves in unmagnetized plasmas.

For uni-directional propagation, Nishikawa *et al* [11] suggested a simpler set whic contains a driven KDV equation instead of the driven Boussinesq equation (15). Note tha the Boussinesq equation has fourth-order space and second-order time derivative whereas the KDV has third-order space and first-order time derivative terms. The latte equation can be systematically and rigorously derived by using the reductive perturbatio analysis [12] on the basic set of plasma equations describing the low-frequency wave However, it can also be directly obtained from the driven Boussinesq equation (15) unde uni-directional, near-sonic approximation which allows us to use $\partial/\partial t \approx -\partial/\partial x$ fo propagation along the positive x-direction. Equation (15) then reduces to the equation,

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}\right)\phi+\frac{1}{2}\frac{\partial^3\phi}{\partial x^3}+\phi\frac{\partial\phi}{\partial x}=-\frac{1}{2}\frac{\partial}{\partial x}\left(\frac{EE^*}{4}\right), \tag{18}$$

which is the driven KDV equation. This equation is coupled, as in the case of th Boussinesq equation, to the Schrödinger-like equation (14). For stationary solutions, th Schrödinger–KDV set yields the equations

$$3\frac{\mathrm{d}^2E_a}{\mathrm{d}\xi^2}=\lambda E_a+\phi E_a, \tag{19}$$

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\xi^2}=-2(1-M)\phi-\phi^2-\frac{E_a^2}{4}, \tag{20}$$

which is very similar to the set of equations (16) and (17). Note that eq. (20) can b directly obtained, as expected, from eq. (17) by using $1-M^2\approx 2(1-M)$ which is vali for $M\to 1$.

3.4 *Exact analytical solutions*

The coupled equations (19) and (20) can be solved for exact analytical solution satisfying localized boundary conditions. The equations have two free parameters, and is desirable to obtain solutions valid in the entire allowed regions of the parameter spac However, it has been possible so far only to obtain exact solutions valid on a straight lin defined by the equation, $M=1-20\Delta/3$ in the two-dimensional space spanned by th parameters M and Δ. The solutions are given [11] explicitly by

$$E_a(\xi)=\pm 16\sqrt{3}\Delta\,\mathrm{sech}(\mu\xi)\tanh(\mu\xi), \tag{21}$$

$$\phi(\xi)=-12\Delta\,\mathrm{sech}^2(\mu\xi), \tag{22}$$

where $\mu^2=2\Delta/3$, and $\lambda\approx 2\Delta$ has been used. Similar solutions also exist [10] for th coupled equations (16) and (17).

It is interesting to compare the exact solution of the Zakharov set with that of th Schrödinger–KDV (or, –Boussinesq) set obtained above. The former has two fr parameters whereas the latter has only one free parameter. In the Zakharov case, th solutions for both the field variables are symmetric with respect to $\xi=0$. On the oth

hand, for the Schrödinger–KDV (or, –Boussinesq) set, the E-field solution is anti-symmetric with respect to $\xi = 0$ whereas the ϕ-field is symmetric as earlier. Clearly, in both cases, the solutions for $E_a^2(\xi)$ have symmetric structures but with different shapes: For the Zakharov set, $E_a^2(\xi)$ is bell-shaped with only a single-hump whereas for the Schrödinger–KDV (or, –Boussinesq) set, it has a double-hump structure with the local minimum at the centre ($\xi = 0$) touching zero. Both the solutions are localized, that is, the fields as well as their derivatives tend to zero as $|\xi| \to 0$. In contrast to eq. (13), the solutions (21) and (22) yield the relation,

$$\frac{E_a^2}{4} = (-16\Delta)\phi + (-\tfrac{4}{3})\phi^2. \tag{23}$$

Thus, the relative scaling of E_a and ϕ are different for the two cases.

3.5 *Exact nonlinear equations*

The equations describing the low-frequency dynamics in the above models have been derived perturbatively using certain approximations. For the Zakharov set, it is a linear wave equation which is derived under the assumption of quasi-neutrality. On the other hand, both the Boussinesq as well as the KDV equations take into account the effects due to charge separation effects only perturbatively. Both are derived under the assumption of weak nonlinearity and, hence, are valid only for small amplitudes. For large amplitude waves, one needs to take into account full nonlinearity as well as charge separation effects by using the Poisson equation. By starting with the relevant fluid equations for the low-frequency dynamics (namely, the continuity and the momentum equations for the ions, and the Boltzmann distribution for the electrons, together with the full Poisson equation), we derived [13, 14] the following set of exact stationary nonlinear equations for the coupled Langmuir and ion-acoustic waves:

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\xi^2} = -\frac{M}{\sqrt{M^2 - 2\phi}} + \exp\left(\phi - \frac{E_a^2}{4}\right), \tag{24}$$

$$3\frac{\mathrm{d}^2 E_a}{\mathrm{d}\xi^2} = (\lambda - 1)E_a + E_a \exp\left(\phi - \frac{E_a^2}{4}\right). \tag{25}$$

While the first equation is essentially the stationary form of the Poisson equation, the second equation arises from the Schrödinger-like equation (8). These equations incorporate exactly all the nonlinearities in the low-frequency dynamics.

It is easy to verify that eq. (24) contains, as limiting cases, the low-frequency equations used in the earlier models. For $|\phi|, E_a^2 \ll 1$, we expand the right-hand side of eq. (24) and keep only the most dominant nonlinear terms to obtain,

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\xi^2} \approx -\frac{(1 - M^2)}{M^2}\phi - \frac{(3 - M^4)}{2M^4}\phi^2 - \frac{E_a^2}{4}. \tag{26}$$

The Zakharov case is obtained by neglecting (because of the quasi-neutrality assumption) the left-hand side of eq. (26) and dropping the nonlinear term in ϕ. This yields

$$\frac{E_a^2}{4} = -\frac{(1 - M^2)}{M^2}\phi, \tag{27}$$

which is just eq. (13). On the other hand, the Boussinesq limit corresponds to taking the near-sonic limit $M^2 \to 1$ in eq. (26) which becomes

$$\frac{d^2\phi}{d\xi^2} = -(1-M^2)\phi - \phi^2 - \frac{E_a^2}{4}, \tag{28}$$

which is same as eq. (17). The KDV limit is trivially obtained from eq. (28), as earlier, by writing $1 - M^2 \approx 2(1-M)$ which is valid for $M \to 1$.

3.5.1 *A method of solution.* The coupled equations (24) and (25) are highly nonlinear and, as such, their exact analytical solutions have not been obtained so far. However, it is possible to find approximate analytical solutions by following a novel method [13, 14]. To this end, we first note that eqs (24) and (25) can be derived from a Lagrangian

$$L = \left\{\frac{1}{2}\left(\frac{d\phi}{d\xi}\right)^2 - \frac{3}{4}\left(\frac{dE_a}{d\xi}\right)^2\right\} - V(E_a, \phi), \tag{29}$$

where the 'potential' $V(E_a, \phi)$ is given by

$$V(E_a, \phi) = -\left[M(M^2 - 2\phi)^{1/2} + \frac{1}{4}(1-\lambda)E_a^2 + \exp\left(\phi - \frac{E_a^2}{4}\right)\right]. \tag{30}$$

Note that the 'kinetic energy' in eq. (29) is not positive definite, but can change sign as ξ varies. The corresponding Hamiltonian is given by

$$H = \left\{\frac{1}{2}\left(\frac{d\phi}{d\xi}\right)^2 - \frac{3}{4}\left(\frac{dE_a}{d\xi}\right)^2\right\} + V(E_a, \phi). \tag{31}$$

Since the Lagrangian L is independent of ξ explicitly, the Hamiltonian (H) is an 'integral of motion', and has a value equal to $-(1+M^2)$ for the localized solutions.

Using the Hamiltonian function H, it is possible to eliminate the independent variable ξ in the two equations (24) and (25) to yield the following equation for $\psi \equiv E_a^2/4$ in terms of ϕ alone:

$$\begin{aligned}
&12\psi[M(M^2-2\phi)^{1/2} + (1-\lambda)\psi + \exp(\phi-\psi) - (1+M^2)]\left(\frac{d^2\psi}{d\phi^2}\right) \\
&\quad + 9[M(M^2-2\phi)^{-1/2} - \exp(\phi-\psi)]\left(\frac{d\psi}{d\phi}\right)^3 - 6[M(M^2-2\phi)^{1/2} \\
&\quad + 2(1-\lambda)\psi + (1-\psi)\exp(\phi-\psi) - (1+M^2)]\left(\frac{d\psi}{d\phi}\right)^2 \\
&\quad - 6\psi[M(M^2-2\phi)^{-1/2} - \exp(\phi-\psi)]\left(\frac{d\psi}{d\phi}\right) \\
&\quad + 4\psi^2[(1-\lambda) - \exp(\phi-\psi)] = 0.
\end{aligned} \tag{32}$$

Equation (32) does not *explicitly* contain the independent variable ξ, and may be considered as the 'trajectory equation' in the (ψ, ϕ) phase plane for the motion of a classical particle in a two-dimensional potential defined by $V(\psi, \phi)$. Any exact solution of

the trajectory equation (32) together with the Hamiltonian (31) yields, in principle, a solution of the coupled equations (24) and (25). However, in the absence of any known exact solution of eq. (32), we look for solutions of the form

$$\psi = \sum_{n=0}^{\infty} a_n \phi^n, \tag{33}$$

where the coefficients a_n which are functions of the free parameters have to be determined suitably. In writing eq. (33), we are guided by the relations (13) and (23) which are obtained from the exact analytical solutions of the approximate equations. For convenience, we introduce the variable, $\theta \equiv \phi/M^2$ so that eq. (33) becomes

$$\psi = \sum_{n=0}^{\infty} b_n \theta^n, \tag{34}$$

where $b_n = M^{2n} a_n$.

Using eq. (34) in eq. (32), and by equating the coefficients of like powers of θ to zero, we are able to determine *uniquely* the coefficients in terms of first order algebraic equations for every n. For localized solutions the first coefficient b_0 is zero whereas the next four coefficients are explicitly given in ref. [14].

To obtain explicit solutions which are of interest for the present discussion, we consider a two-term approximation to the expansion in eq. (34) and write

$$\psi = b_1 \theta + b_2 \theta^2. \tag{35}$$

Using this expression in eq. (24) and expanding the right-hand side, we obtain, after retaining terms up to ϕ^3, the following equation for $\theta(\xi)$:

$$M^2 \frac{\mathrm{d}^2\theta}{\mathrm{d}\xi^2} = \alpha_1 \theta + \alpha_2 \theta^2 + \alpha_3 \theta^3, \tag{36}$$

where the coefficients α_1, α_2 and α_3 are known functions of b_1 and b_2 [13, 14]. The localized solution of eq. (36) is given by

$$\theta(\xi) = \frac{\beta_1 \beta_2 \operatorname{sech}^2(\mu\xi)}{\beta_1 - \beta_2 \tanh^2(\mu\xi)}, \tag{37}$$

where $\mu^2 = \lambda/3$, and β_1 and β_2 are known functions of α_1, α_2 and α_3. The solution for $E_a(\xi)$ is obtained by simply substituting the solution (37) into eq. (35), that is

$$E_a^2(\xi) = 4[b_1 \theta(\xi) + b_2 \theta^2(\xi)]. \tag{38}$$

This completes the approximate solution of the coupled eqs (24) and (25).

Equations (37) and (38) can be analysed [14] for the existence of localized solutions in the (M, Δ) parameter space. For a given $\Delta < 1$, the solutions $E^2(\xi)$ and $\phi(\xi)$ have, respectively, single-hump and single-dip structures for sufficiently small Mach numbers ($M \ll 1$). This corresponds to the case of the solutions obtained from the Zakharov equations. When the Mach number is increased, the solution for $E^2(\xi)$ flattens at the top till a critical Mach number M_{crit} is reached. For further increase in the Mach number, $E^2(\xi)$ develops a local dip around the center $\xi = 0$. The depth of the dip increases with

increase in the Mach number, and $E^2(\xi = 0)$ becomes zero when M is equal to a cut-off Mach number M_{cut}. Beyond this value of the Mach number, $E^2(\xi)$ becomes negative around $\xi = 0$ which violates the boundary conditions, and hence the solutions are not valid for $M > M_{cut}$. The solutions on the line $M = M_{cut}$ correspond to the solutions obtained from the Schrödinger–Boussinesq (or, –KDV) equations. In fact, for $M = M_{cut}$ and for sufficiently small values of Δ, the explicit solutions (37) and (38) can be shown to exactly reduce to the solutions (21) and (22) [anti-symmetric $E(\xi)$ and symmetric $\phi(\xi)$] obtained from the Schrödinger–Boussinesq (or, –KDV) case. For all values of M, the solution for $\phi(\xi)$ has always a single-dip, symmetric structure whose amplitude increases with the increase in the Mach number reaching the maximum value at $M = M_{cut}$.

Thus, the approximate solutions (37) and (38) of the exact governing equations (24) and (25) exactly reduce, for near-sonic propagations, to the exact solutions (21) and (22) of the approximate governing equations (19) and (20). It should be noted that there is no reason why such a complete equivalence between the two should exist at all since the approximations involved in the solutions as well as the equations are entirely of different nature. On the other hand, the very existence of such an exact reduction lends support to the suitability of the above method of solution used in solving coupled equations.

The coupled Schrödinger–Boussinesq (or, –KDV) system of equations occurs in many different problems in plasma physics where a high-frequency wave is coupled to a suitable low-frequency wave via the ponderomotive force. It has been shown that the coupled electromagnetic and ion-acoustic waves [15, 16] as well as the upper-hybrid and the magnetoacoustic waves [17] are indeed governed by a Schrödinger–Boussinesq (or, –KDV) system but with different sets of free parameters. In fact, the latter system with arbitrary free parameters for each of the terms in the two equations constitutes a general set which for stationary solutions yields a generic system of two coupled ordinary differential equations admitting different classes of exact analytical solutions. This is discussed in the next section.

4. Generic system of equations

Consider a Schrödinger-like equation with arbitrary free parameters in the form

$$i\left(\frac{\partial E}{\partial t} + \lambda_1 \frac{\partial E}{\partial x}\right) + \lambda_2 \frac{\partial^2 E}{\partial x^2} = \lambda_3 E + \lambda_4 \phi E, \tag{39}$$

which is coupled to the Boussinesq equation

$$\left(\frac{\partial^2}{\partial t^2} + \mu_1 \frac{\partial^2}{\partial x^2}\right)\phi + \mu_2 \frac{\partial^4 \phi}{\partial x^4} + \mu_3 \frac{\partial^2}{\partial x^2}(\phi^2) = \mu_4 \frac{\partial^2}{\partial x^2}(EE^*), \tag{40}$$

or, to the KDV equation

$$\left(\frac{\partial}{\partial t} + \mu_1 \frac{\partial}{\partial x}\right)\phi + \mu_2 \frac{\partial^3 \phi}{\partial x^3} + \mu_3 \phi \frac{\partial \phi}{\partial x} = \mu_4 \frac{\partial}{\partial x}(EE^*), \tag{41}$$

where λ_i and μ_i, $i = 1, 2, 3, 4$ are arbitrary free parameters which can take different values depending on the problem at hand. In a stationary frame $\xi = x - Mt$, the above equations

$$\beta \frac{d^2E}{d\xi^2} = b_1 E + b_2 \phi E, \tag{42}$$

$$\lambda \frac{d^2\phi}{d\xi^2} = d_1 \phi + d_2 \phi^2 + d_3 E^2, \tag{43}$$

where $\lambda, \beta, b_1, b_2, d_1, d_2$ and d_3 are the free parameters defined in terms of λ_i and μ_i, $i = 1, 2, 3, 4$.

Equations (42) and (43) constitute a generic set of equations having seven free parameters. However, by a proper rescaling of the variables, it is possible to reduce the number of free parameters. This will be considered in a later section [cf. eqs (65) and (66)]. It is of interest to find exact analytical solutions of the generic equations valid in as much of the region in the parameter space as possible. This is done by following the method of solution described in § 3.5.1. As earlier, we try a series solution of the form given by eq. (33) which, in general, does not truncate. However, by properly choosing certain curves or surfaces in the parameter space, it is possible to make the coefficients b_n become zero for n greater than a certain value, say, m. The resulting polynomial relationship between E and ϕ thus becomes an *exact solution* of the corresponding 'trajectory equation'. This relation together with the corresponding Hamiltonian function for the generic system yield different classes of the exact analytical solutions of the eqs (42) and (43). Omitting the details which can be found in ref. [18], we summarize below the various classes of explicit solutions.

4.1 *Exact analytical solutions*

The following classes of exact analytical solutions of the generic system of equations (42) and (43) have been obtained [18] so far:

(A) For $\beta d_1 b_2 - 2b_1(3\beta d_2 - \lambda b_2) = 0$:

$$E(\xi) = \pm \frac{6b_1}{b_2} \left(\frac{\beta d_2 - \lambda b_2}{\beta d_3} \right)^{1/2} \operatorname{sech}(\mu\xi) \tanh(\mu\xi), \tag{44}$$

$$\phi(\xi) = -\frac{6b_1}{b_2} \operatorname{sech}^2(\mu\xi); \quad \mu = \left(\frac{b_1}{\beta} \right)^{1/2}. \tag{45}$$

(B) For $3\lambda b_2 - \beta d_2 = 0$:

$$E(\xi) = \pm \left[\frac{6\lambda b_1}{\beta^2 d_2 d_3} (\beta d_1 - 4\lambda b_1) \right]^{1/2} \operatorname{sech}(\mu\xi), \tag{46}$$

$$\phi(\xi) = -\frac{2b_1}{b_2} \operatorname{sech}^2(\mu\xi); \quad \mu = \left(\frac{b_1}{\beta} \right)^{1/2}. \tag{47}$$

(C) For $\beta d_1 + 2\lambda b_1 = 0$ and $d_2 = 0$: [In this case, eq. (43) is linear in ϕ.]

$$E(\xi) = \pm \left(\frac{18 b_1 d_1}{d_3 b_2} \right)^{1/2} \operatorname{sech}(\mu\xi) \tanh(\mu\xi), \tag{48}$$

$$\phi(\xi) = \frac{3\beta d_1}{\lambda b_2}\operatorname{sech}^2(\mu\xi); \quad \mu = \left(\frac{b_1}{\beta}\right)^{1/2}. \tag{49}$$

(D) For $\lambda = 0$: [In this case, eq. (43) is singular and yields just an algebraic relation between E and ϕ.]

For this value of λ, there are two cases:

(i) For $b_2 d_1 - 6 b_1 d_2 = 0$,

$$E(\xi) = \pm \frac{d_1}{(d_2 d_3)^{1/2}} \operatorname{sech}(\mu\xi)\tanh(\mu\xi), \tag{50}$$

$$\phi(\xi) = -\frac{d_1}{d_2}\operatorname{sech}^2\left[\left(\frac{b_1}{\beta}\right)^{1/2}\xi\right]; \quad \mu = \left(\frac{d_1 b_2 - 2 d_2 b_1}{4\beta d_2}\right)^{1/2}. \tag{51}$$

(ii) For $5 b_2 d_1 - 6 b_1 d_2 = 0$,

$$E(\xi) = \pm \frac{d_1}{(d_2 d_3)^{1/2}} \operatorname{sech}(\mu\xi)\tanh(\mu\xi), \tag{52}$$

$$\phi(\xi) = -\frac{d_1}{d_2}\tanh^2\left\{\left(\frac{-b_1}{5\beta}\right)^{1/2}\xi\right\}; \quad \mu = \left(\frac{d_1 b_2 - 2 d_2 b_1}{4\beta d_2}\right)^{1/2}. \tag{53}$$

(E) Periodic solutions: In addition to the above solutions, a new class of periodic solutions were recently reported in ref. [3]. These were found by trying a solution of the form $\phi = C_0 + C_1 E$ for the 'complementary trajectory equation' for ϕ in terms of E. Two sets of parameter values were found to be admissible:

(i) For $C_0 = 0$, one requires, $\lambda b_1 - \beta d_1 = 0$;
(ii) For $d_1 + d_2 C_0 = 0$ one requires $\lambda b_1 + C_0(\lambda b_2 - \beta d_2) = 0$.

In both cases, C_1 is given by

$$C_1^2 = \frac{\beta d_3}{\lambda b_2 - \beta d_2}. \tag{54}$$

The corresponding explicit solutions are given by

$$E(\xi) = h_1 sn^2(\mu\xi, k) + h_2 cn^2(\mu\xi, k), \tag{55}$$

$$\phi(\xi) = C_0 + C_1 E(\xi), \tag{56}$$

where

$$\mu^2 = \frac{b_2 C_1}{6\beta}(h_3 - h_2), \quad k^2 = \frac{h_1 - h_2}{h_3 - h_2} \tag{57}$$

and h_1, h_2 and h_3 are the three real roots of the cubic equation

$$h^3 + \frac{3(b_1 + b_2 C_0)}{2 b_2 C_1} h^2 + \frac{3C}{2 b_2 C_1} = 0. \tag{58}$$

In eq. (55), sn and cn are the Jacobian elliptic functions, and, in eq. (58), C is a constant of integration. The solutions are, in general, periodic but for $C = 0$ are localized and

given by

$$E(\xi) = -\frac{3(b_1 + b_2 C_0)}{2b_2 C_1} \operatorname{sech}^2(\mu\xi), \tag{59}$$

$$\phi(\xi) = C_0 + C_1 E(\xi); \quad \mu^2 = \frac{b_1 + b_2 C_0}{4\beta}. \tag{60}$$

Note that the solution for $\phi(\xi)$ is truly localized only when $C_0 = 0$.

5. Relation to other systems

The generic equations (42) and (43) have relevance to some other equations that are commonly used in other branches of physics. We consider below some particular examples.

5.1 *Hénon–Heiles system*

The Hénon–Heiles system has been extensively studied in the field of nonlinear dynamics since it was first proposed by Hénon and Heiles [4]. The generalized form of these equations with arbitrary coefficients can be obtained from the Hamiltonian [19]

$$H = \tfrac{1}{2}(p_x^2 + p_y^2) + \tfrac{1}{2}(Ax^2 + By^2) + (Dx^2 y - \tfrac{1}{3}Cy^3), \tag{61}$$

where A, B, C and D are arbitrary real parameters, x and y are the spatial coordinates, and p_x and p_y are the corresponding conjugate momenta. Clearly, H is simply the Hamiltonian for a two-dimensional harmonic oscillator with certain specific nonlinear terms given by the last two terms in eq. (61). The standard form first investigated by Hénon and Heiles [4] corresponds to the case when $A = B = C = D = 1$. The fundamental question of general interest in such Hamiltonian systems is to identify the parameters for which the system is 'completely integrable'. Since the system is of two-degrees of freedom and the Hamiltonian is an 'integral of motion', the problem is equivalent to finding another integral of motion that is in involution with the Hamiltonian.

By proper normalization of the variables, the Hamiltonian (61) can be re-written in the canonical form, namely

$$H = \tfrac{1}{2}(p_x^2 + p_y^2) + \tfrac{1}{2}(-p_1 x^2 - d_1 y^2) + (x^2 y + \tfrac{1}{3}d_2 y^3), \tag{62}$$

where p_1 takes values $+1$ or -1, and d_1 and d_2 are the two free parameters. The corresponding equations of motion are given by

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = p_1 x - 2xy, \tag{63}$$

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} = d_1 y - d_2 y^2 - x^2. \tag{64}$$

These equations are known to be integrable [20] for the following three sets of parameter values:

(C) $a_1 = 10p_1, a_2 = +10$.

Note that in all the cases d_2 is always positive which indicates that for the known integrable cases, the nonlinear terms in the equations of motion, namely, eqs (63) and (64) have the same sign.

Consider now the generic equations (42) and (43) whose variables can be suitably normalized to yield the canonical set

$$\frac{d^2E}{d\xi^2} = p_1E - 2\phi E, \tag{65}$$

$$\frac{d^2\phi}{d\xi^2} = d_1\phi - d_2\phi^2 - p_2E^2, \tag{66}$$

where p_2 can take values $+1$ or -1. The Hamiltonian for the eqs (65) and (66) is given by

$$H = \frac{1}{2}\left[p_2\left(\frac{dE}{d\xi}\right)^2 + \left(\frac{d\phi}{d\xi}\right)^2\right] + \frac{1}{2}(-p_1p_2E^2 - d_1\phi^2) + \left(p_2E^2\phi + \frac{1}{3}d_2\phi^3\right). \tag{67}$$

Comparing the two Hamiltonians given by eqs (62) and (67) [or, the equations of motion, (63), (64) and (65), (66)] we note that they are structurally very similar. In fact, for the case when $p_2 = +1$, they are exactly same if we identify the set (E, ϕ) with (x, y). However, as pointed out earlier, for all modulational problems, one finds $p_2 = -1$ and hence the kinetic energy in (67) is not positive definite. Thus, the stationary equations governing the nonlinear development of the modulational instability of an high-frequency wave coupled to a suitable low-frequency wave in plasmas are complementary to the Hénon–Heiles equations, but seem to have fundamentally different qualitative features. For example, in the case of the usual Hénon–Heiles system (with positive definite kinetic energy), any minimum in the potential guarantees local oscillatory motion of the particle. This need not be true of the generic equations since the 'kinetic energy', term can change its sign during the motion of the 'particle'. The exact solutions obtained in §4.1 do not, however, guarantee the 'complete integrability' of the equations since they satisfy specialized boundary conditions. On the other hand, even such specialized solutions valid in the entire allowed regions of the parameter space have also not been obtained so far.

The integrable parameter regimes for the generic Hamiltonian (H) given by eq. (67) for $p_2 = +1$ or -1 has recently been discussed elsewhere [21]. In particular, the generic Hamiltonian is integrable for three sets of parameter values and it is possible to explicitly obtain, in each case, the corresponding second integral of motion. Results show that the integrable parameter regimes in the two cases corresponding to $p_2 = +1$ and -1 are complementary to each other. Furthermore, there exists a direct one-to-one correspondence between the known integrable cases of H and the stationary Hamiltonian flows associated with the only integrable nonlinear evolution equations (of polynomial and autonomous type) with a scale-weight of seven. The latter result strongly suggests that there are possibly no other integrable cases of H for both $p_2 = +1$ as well as $p_2 = -1$. An

extended discussion on some of these questions together with the relevant details on the known results can be found in ref. [21].

5.2 *Self-dual Yang–Mills system*

In the last couple of years, there has been much interest in the classical Yang–Mills field equations, particularly in their reduction to simpler nonlinear equations [5] by using certain symmetry properties. The classical Yang–Mills equations are written in the form

$$D_\mu G_{\mu\nu} \equiv \partial_\mu G_{\mu\nu} + [A_\mu, G_{\mu\nu}] = 0, \tag{68}$$

where

$$G_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu], \tag{69}$$

where $[A_\mu, A_\nu]$ is a suitable Lie bracket defined over the Yang–Mills field variables (A_μ), and the subscripts for ∂ denote the partial derivatives. The Yang–Mills fields are said to be self-dual if the condition

$$\tilde{G}_{\mu\nu} \equiv \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma} G_{\rho\sigma} = +G_{\mu\nu}, \tag{70}$$

is satisfied where $\epsilon_{\mu\nu\rho\sigma}$ is the usual anti-symmetric tensor. Solutions satisfying the above condition satisfy also the Yang–Mills field equations. Using the gauge degree of freedom implied by the self-dual condition, it has been recently shown [5] that the Yang–Mills field equations can be reduced either to the nonlinear Schrödinger or to the KDV equation. Since nonlinear Schrödinger equation can generally be thought of as a special case (namely, the static limit) of the Zakharov or the Schrödinger–Boussinesq (or, –KDV) system, one would expect to reduce the self-dual Yang–Mills system to the Schrödinger–Boussinesq (or, –KDV) system. Since the latter system is known to yield, for stationary solutions, the generalized Hénon–Heiles system, this would establish a cascading connection from the Yang–Mills to the Hénon–Heiles system via the Schrödinger–Boussinesq (or, –KDV) system. It would also provide a good model to study the 'nonlinear dynamical' behaviour of classical Yang–Mills fields. While such a possibility is quite exciting, even a formulation of the problem is yet to be carried out.

5.3 *Complex KDV equation*

The peculiar nature of the Hamiltonian (67) for the generic system of equations (65) and (66) admitting the case when the 'kinetic energy' term is indefinite is exhibited also by the usual KDV equation when the dependent variable is made complex [6, 7]. Consider the KDV equation

$$\frac{\partial u}{\partial t} + \alpha u \frac{\partial u}{\partial x} + \beta \frac{\partial^3 u}{\partial x^3} = 0, \tag{71}$$

where α and β are the free parameters, and all the variables are real quantities. For the stationary solutions depending on a single variable $\xi \equiv x - Mt$ with one free parameter M, eq. (71) yields

$$\beta \frac{\mathrm{d}^2 u}{\mathrm{d}\xi^2} = Mu - \tfrac{1}{2}\alpha u^2. \tag{72}$$

The Hamiltonian for the above equation is

$$H = \frac{1}{2}\beta\left(\frac{du}{d\xi}\right)^2 + V(u), \tag{73}$$

where the potential $V(u)$ is given by

$$V(u) = -\tfrac{1}{2}Mu^2 + \tfrac{1}{6}\alpha u^3. \tag{74}$$

We now 'complexify' the KDV equation by making the dependent variable u complex and write, $u = u_1 + iu_2$. Equation (72) then yields the set of equations

$$\beta\frac{d^2u_1}{d\xi^2} = Mu_1 - \tfrac{1}{2}\alpha u_1^2 + \tfrac{1}{2}\alpha u_2^2, \tag{75}$$

$$\beta\frac{d^2u_2}{d\xi^2} = Mu_2 - \alpha u_1 u_2. \tag{76}$$

Accordingly, the potential becomes, $V(u) \to V(u_1, u_2) = V_1(u_1, u_2) + iV_2(u_1, u_2)$ where

$$V_1 = -\tfrac{1}{2}M(u_1^2 - u_2^2) + \tfrac{1}{6}\alpha(u_1^3 - 3u_1u_2^2), \tag{77}$$

$$V_2 = -Mu_1u_2 - \tfrac{1}{6}\alpha(u_2^3 - 3u_1^2u_2), \tag{78}$$

and the Hamiltonian $H \to H_1 + iH_2$ where

$$H_1 = \frac{1}{2}\beta\left[\left(\frac{du_1}{d\xi}\right)^2 - \left(\frac{du_2}{d\xi}\right)^2\right] + V_1(u_1, u_2), \tag{79}$$

$$H_2 = \beta\left(\frac{du_1}{d\xi}\right)\left(\frac{du_2}{d\xi}\right) + V_2(u_1, u_2). \tag{80}$$

By identifying u_1 with ϕ and u_2 with E, we note that the eqs (75) and (76) are structurally very similar to the generic equations (42) and (43).

Equations (75) and (76) can also be written in the form

$$\beta\frac{d^2u_1}{d\xi^2} = -\frac{\partial V_1}{\partial u_1}, \quad \beta\frac{d^2u_2}{d\xi^2} = +\frac{\partial V_1}{\partial u_2}, \tag{81}$$

which involve only the potential $V_1(u_1, u_2)$ given by eq. (77). Note that unlike the equations of motion for a classical particle with two degrees of freedom, the second equation has a positive sign for the derivative on the right-hand side. This is a reflection of the fact that the kinetic energy is indefinite.

Using the Cauchy–Riemann conditions, namely

$$\frac{\partial V_1}{\partial u_1} = \frac{\partial V_2}{\partial u_2}, \quad \frac{\partial V_1}{\partial u_2} = -\frac{\partial V_2}{\partial u_1}, \tag{82}$$

equations (81) can also be written in terms of the potential $V_2(u_1, u_2)$ given by (78):

$$\beta\frac{d^2u_1}{d\xi^2} = -\frac{\partial V_2}{\partial u_2}, \quad \beta\frac{d^2u_2}{d\xi^2} = -\frac{\partial V_2}{\partial u_1}. \tag{83}$$

While both the equations have now negative signs for the derivatives on the right-hand sides (like in the case of the equations of motion in classical dynamics), the equation for u_1 has derivative of V_2 with respect to u_2, and vice versa for u_2.

5.4 *'Complexified' classical dynamics*

The above features of the 'complexified' KDV equation are also exhibited by equations obtained from classical dynamical systems with one-degree of freedom by making the dependent variable complex [6]. Consider a one-degree of freedom conservative system defined by the Lagrangian

$$L(q,\dot{q}) = \tfrac{1}{2}\dot{q}^2 - V(q), \tag{84}$$

where dot denotes the time derivative, and the corresponding Hamiltonian

$$H(q,p) = \tfrac{1}{2}p^2 + V(q), \tag{85}$$

where the conjugate momentum (p) is defined by $p = \dot{q}$. Clearly, the equation of motion is

$$\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} = -\frac{\partial V}{\partial q}. \tag{86}$$

Let us now make the dependent variable complex, and write, $q = q_1 + iq_2$. Accordingly, the conjugate momentum becomes complex, that is $p \to p_1 + ip_2$. Then, (q_1, p_1) and (q_2, p_2) constitute canonically conjugate variables. Under complexification, the potential $V(q)$ becomes complex, that is, $V(q) \to V(q_1, q_2) = V_1(q_1, q_2) + iV_2(q_1, q_2)$ where V_1 and V_2 satisfy the Cauchy–Riemann condition

$$\frac{\partial V_1}{\partial q_1} = \frac{\partial V_2}{\partial q_2}, \quad \frac{\partial V_1}{\partial q_2} = -\frac{\partial V_2}{\partial q_1}. \tag{87}$$

Similarly, the Lagrangian and the Hamiltonian yield, respectively,

$$L(q) \to L_1 + iL_2 \equiv [\tfrac{1}{2}(\dot{q}_1^2 - \dot{q}_2^2) - V_1(q_1, q_2)] + i[\dot{q}_1\dot{q}_2 - V_2(q_1, q_2)], \tag{88}$$

$$H(q,p) \to H_1 + iH_2 \equiv [\tfrac{1}{2}(p_1^2 - p_2^2) + V_1(q_1, q_2)] + i[p_1 p_2 + V_2(q_1, q_2)]. \tag{89}$$

Note that the kinetic energy term in L_1 and H_1 is not positive definite.

The Newton's equations of motion can be obtained from the usual Euler–Lagrange equation using the Lagrangian L_1. This yields,

$$\frac{\mathrm{d}^2 q_1}{\mathrm{d}t^2} = -\frac{\partial V_1}{\partial q_1}, \quad \frac{\mathrm{d}^2 q_2}{\mathrm{d}t^2} = +\frac{\partial V_1}{\partial q_2}. \tag{90}$$

Note that here only the potential V_1 is involved and the second equation has a positive sign before the derivative on the right-hand side. On the other hand, one can use the Lagrangian L_2 to obtain the equations in terms of V_2 only:

$$\frac{\mathrm{d}^2 q_1}{\mathrm{d}t^2} = -\frac{\partial V_2}{\partial q_2}, \quad \frac{\mathrm{d}^2 q_2}{\mathrm{d}t^2} = -\frac{\partial V_2}{\partial q_1}. \tag{91}$$

Here, both the equations have negative signs before the derivatives on the right-hand sides. However, the equation for q_1 is defined in terms of the derivative of V_2 with respect to q_2, and vice versa for q_2. Similar features were also noticed in the previous section for the complex KDV equation. As expected, the two sets of eqs (90) and (91) are identical in view of the Cauchy–Riemann conditions (87).

The above peculiarities are also reflected in the Hamilton's equations of motion. Using H_1 in the usual Hamilton's canonical equations of motion, we obtain

$$\dot{q}_1 = p_1, \quad \dot{p}_1 = -\frac{\partial V_1}{\partial q_1}, \tag{92}$$

$$\dot{q}_2 = -p_2, \quad \dot{p}_2 = -\frac{\partial V_1}{\partial q_2}. \tag{93}$$

Note that if we eliminate p_1 from the first set and p_2 from the second set of equations, we recover the equations of motion given by (90). On the other hand, if we use, instead, the Hamiltonian H_2, we obtain the equations

$$\dot{q}_1 = p_2, \quad \dot{p}_1 = -\frac{\partial V_2}{\partial q_1}, \tag{94}$$

$$\dot{q}_2 = p_1, \quad \dot{p}_2 = -\frac{\partial V_2}{\partial q_2}. \tag{95}$$

Unlike the previous case, here cross mixing of the equations is necessary in order to eliminate p_1 and p_2 which leads to the equations of motion given by (91).

6. Summary

To summarize, we have discussed the properties as well as some exact analytical solutions of a pair of coupled nonlinear ordinary differential equations that occur in different branches of physics. In particular, in the field of plasma physics, such equations govern the stationary solutions of time-dependent coupled nonlinear equations [like the Schrödinger–Boussinesq (or, –KDV) system] that describe the nonlinear evolution of the amplitude modulated high-frequency waves coupled to suitable low-frequency waves. The latter system with arbitrary free parameters leads to a generic system of coupled scalar field equations whose exact analytical solutions valid in different regions of the parameter space have been explicitly obtained.

The generic system is shown to contain as special cases equations that have been studied in other fields. For example, the case when the kinetic energy in the corresponding Hamiltonian is positive definite is shown to exactly correspond to the well-known 'Hénon–Heiles' system (in the field nonlinear dynamics) in a generalized form. On the other hand, the equations obtained from the stationary KDV equation by making the dependent variable complex are shown to be structurally similar to the generic system. We have also suggested the possibility of relating the classical Yang–Mills equations to the Hénon–Heiles system by reducing them to the Schrödinger–Boussinesq (or, –KDV) system using the self-duality condition. The Hamiltonian of the generic system for the nonlinear wave modulations admits kinetic energy term which is not positive definite. It is shown that this feature is exhibited also by the equations obtained

from the classical equation of motion for a particle with one-degree of freedom in a conservative potential field by making the dependent variable complex.

References

[1] R Rajaram, *Phys. Rev. Lett.* **42**, 200 (1979)
[2] A Hasegawa, *Plasma Instabilities and Nonlinear Effects* (Springer, Berlin, 1975) p. 194
[3] N N Rao and D J Kaup, *J. Phys.* **A24**, L993 (1991)
[4] M Hénon and C Heiles, *Astron. J.* **69**, 73 (1964)
[5] L J Mason and G A J Sparling, *Phys. Lett.* **A137**, 29 (1989)
S Chakravarthy, M J Ablowitz and P A Clarkson, *Phys. Rev. Lett.* **63**, 1085 (1990)
[6] N N Rao, B Buti and S B Khadkikar, *Pramana – J. Phys.* **27**, 497 (1986)
[7] B Buti, N N Rao and S B Khadkikar, *Phys. Scr.* **34**, 729 (1980)
[8] F F Chen, *Introduction to Plasma Physics* (Plenum, New York, 1974) p. 256
[9] V E Zakharov, *Sov. Phys. – JETP* **35**, 908 (1972)
[10] V G Makhankov, *Phys. Lett.* **A50**, 42 (1974)
[11] K Nishikawa, H Hojo, K Mima and H Ikezi, *Phys. Rev. Lett.* **33**, 148 (1974)
[12] H Washimi and T Taniuti, *Phys. Rev. Lett.* **17**, 966 (1966)
R C Davidson, *Methods in Nonlinear Plasma Theory* (Academic, New York, 1972) ch. 2
[13] R K Varma and N N Rao, *Phys. Lett.* **A79**, 311 (1980)
[14] N N Rao and R K Varma, *J. Plasma Phys.* **27**, 95 (1982)
[15] N N Rao, R K Varma, P K Shukla and M Y Yu, *Phys. Fluids* **26**, 2488 (1983)
[16] N N Rao, *Phys. Rev.* **A37**, 4846 (1988)
[17] N N Rao, *J. Plasma Phys.* **39**, 385 (1988)
[18] N N Rao, *J. Phys.* **A22**, 4813 (1989)
[19] Y F Chang, M Tabor and J Weiss, *J. Math. Phys.* **23**, 531 (1982)
[20] M Lakshmanan and R Sahadevan, *Phys. Rep.* **224**, 1 (1993)
[21] N N Rao, *Pramana – J. Phys.* **46**, 161 (1996)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
p. 129

Thermodynamics of nonlinear plasma systems

P K KAW
Institute for Plasma Research, Bhat, Gandhinagar 382 424, India

Full text of this article has appeared in *Current Science* Vol. 59, Nos 21 & 22, pp. 1065–1079, 25 November 1990.

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 131–145

The large scale structure of the universe: Dynamical and statistical aspects

VARUN SAHNI
Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkind, Pune 411 007, India

Abstract. The description of gravitational clustering is essentially statistical but its origin is dynamical. Hence both aspects of clustering: dynamical and statistical, must be understood in order to arrive at a proper appreciation of the subject of gravitational instability and the formation and evolution of the large scale structure of the Universe. Key dynamical aspects of gravitational clustering such as the Zeldovich approximation and its extension – the adhesion model are reviewed. Statistical indicators of clustering such as the correlation function and percolation theory, as applied to the large scale structure of the Universe have also been focussed on.

Keywords. Large scale structure; cosmology; correlation function; percolation theory.

PACS No. 95.30

1. Introduction

Observations of the large scale structure of the universe can be conveniently grouped under three main categories: (i) those that deal with the clustering properties of galaxies (these include statistical indicators such as: correlation functions, counts in cells, fractal analysis etc.); (ii) the peculiar velocities of galaxies after the smooth Hubble flow has been subtracted out; and (iii) the anisotropy of the cosmic microwave background radiation (CMB). (The latter provides a probe of the amplitude and spectrum of density inhomogeneities in the distant past, soon after the cosmological recombination of hydrogen.)

I shall briefly dwell on each of these observational aspects of large scale structure before proceeding to discuss related dynamical and statistical aspects of the formation and evolution of large scale structure in the universe.

1.1 *Clustering of galaxies and clusters*

The fact that galaxies are not randomly distributed on the celestial sphere can be easily discerned from projected two-dimensional catalogues. (Over 2 million galaxies have been mapped in 2D surveys, although the redshifts of only a small fraction of these is accurately known at present.) The clustering of galaxies is amply borne out in three dimensional redshift survey's such as the CfA survey and the southern sky redshift survey (SSRS), which, taken together contain over 15,000 galaxies. As figure 1 demonstrates,

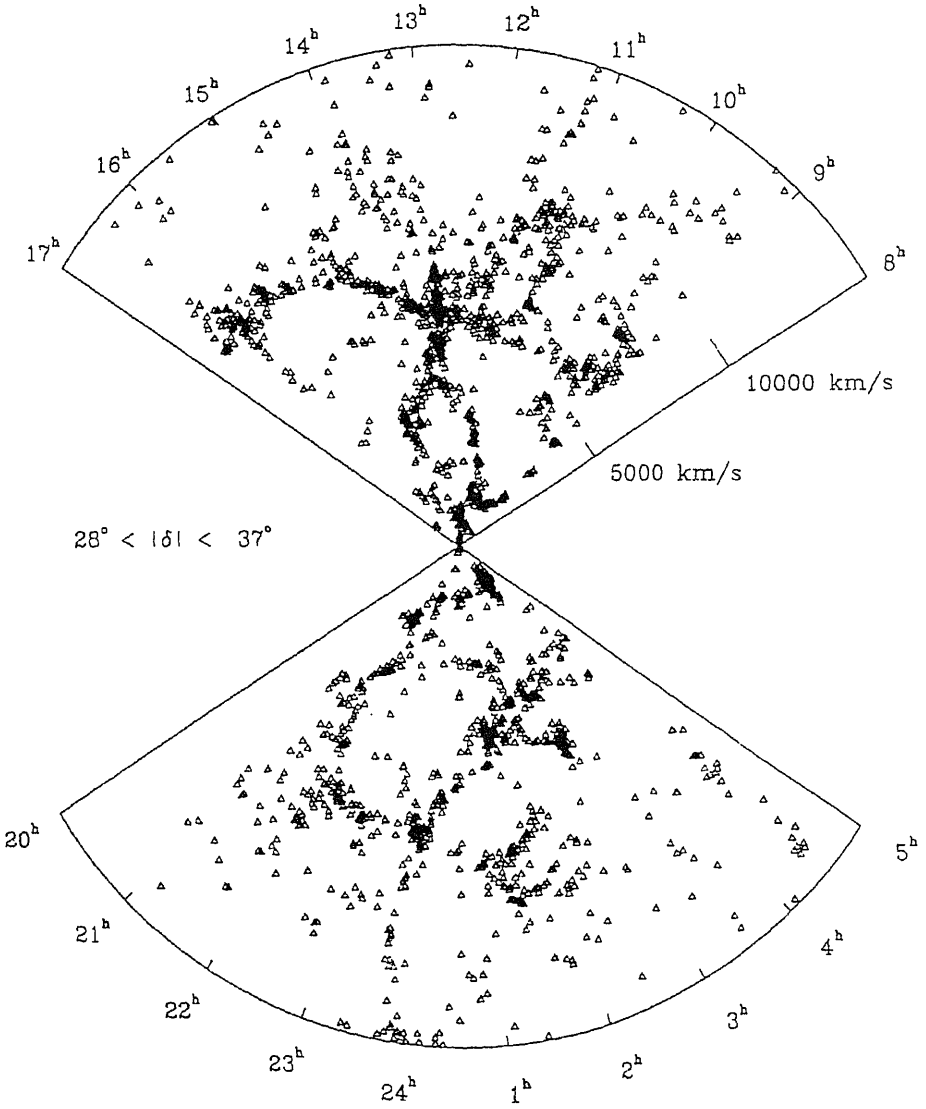


Figure 1. The right ascension is plotted against the radial velocity field for galaxies in the SSRS2-CfA2 redshift surveys. The combined surveys contain over 15,000 galaxies of which a fraction are shown in this figure [28]. The northern survey shows the coma cluster and the great wall. Several wall-like features and large voids are also present in the Southern survey.

galaxies appear to lie preferentially along sheets and filaments which are separated from each other by vast empty regions called voids. The typical length scale characterizing cellular structure in the universe is $\sim 30 - 50\,h^{-1}$ Mpc., h being the Hubble parameter in units of 100 km/sec/Mpc.

The simplest indicator of clustering is the two point correlation function $\xi(r)$, which when applied to galaxies estimates the probability (in excess of random), of finding a galaxy at a distance r from a given galaxy. (For a uniform Poisson process $\xi(r) = 0$.) On scales $r < 10\,h^{-1}$ Mpc the two point correlation function can be accurately described by a power-law [1]

$$\xi_{gg}(r) = \left(\frac{r}{r_{0,g}}\right)^{-1.8} \qquad r_{0,g} \simeq 5\,h^{-1}\,\text{Mpc}. \tag{1}$$

It is useful to express the density contrast in the universe as a Fourier transform

$$\delta(\mathbf{x},t) = \frac{\rho - \bar{\rho}}{\bar{\rho}} = \int \mathrm{d}^3k e^{i\mathbf{k}\cdot\mathbf{x}} \hat{\delta}(\mathbf{k},t). \tag{2}$$

$\hat{\delta}(\mathbf{k},t) = |\delta_k(t)| \exp(i\phi_k)$ (For a gaussian random field (GRF) the amplitudes δ_k have a Rayleigh distribution and the phases ϕ_k are independent and uniformly random on the interval $[0, 2\pi]$.) A GRF is completely specified by its power spectrum $P(k) = \langle|\delta_k|^2\rangle$.

The correlation function is an important measure of clustering since $\xi(r)$ and $P(k)$ – the power spectrum of density fluctuations, form a Fourier transform pair

$$\xi(r) = \int \mathrm{d}^3\mathbf{k} \exp(i\mathbf{k}\cdot\mathbf{r}) P(k), \tag{3}$$

$$P(k) = \frac{1}{(2\pi)^3} \int \mathrm{d}^3\mathbf{x} \exp(-i\mathbf{k}\cdot\mathbf{x}) \xi(x). \tag{4}$$

In standard models of structure formation based on gravitational instability the power spectrum $P(k)$ is related to (i) the spectrum of primordial fluctuations generated during inflation, and (ii) the nature of dark matter (whether 'hot' or 'cold'). The power spectrum in models with 'hot dark matter' shows a sharp cutoff on scales ≤ 40 Mpc., (related to the free-streaming distance of the 'hot' light dark matter particle) so that super-cluster size objects are the first to form in such scenarios. Galaxies form later due to the fragmentation of primordial super-pancakes. This scenario is commonly referred to as the 'top-down' scenario for galaxy formation, and was first proposed by Zeldovich and co-workers in Moscow during the 1970s and 80s in connection with adiabatic baryonic models of structure formation.

The cold dark matter model (CDM) on the other hand, has power on all scales which is mainly related to the fact that the dark matter particle in this case is assumed to be 'cold' – having negligible free-streaming distance. As a result the first objects to form in this scenario have a low mass, and larger mass objects (galaxies and clusters) form due to the hierachical merging of smaller mass objects, leading to a 'bottom-up' scenario of structure formation.

It is also interesting that clusters of galaxies (which typically have $\geqslant 100$ members), have a correlation function which is similar in form to equation (1), but with a much larger correlation length: $r_{0,c} \simeq 12 - 25\,h^{-1}$ Mpc. This trend is followed by richer (and rarer) objects which seem to display stronger clustering as shown in figure 2 where the amplitude of the two-point correlation function (scaled to 1 Mpc) between objects belonging to a given richness class (galaxy groups, Abell clusters, X-ray clusters, super-clusters etc.) is shown as a function of their mean separation. The reason for the increase

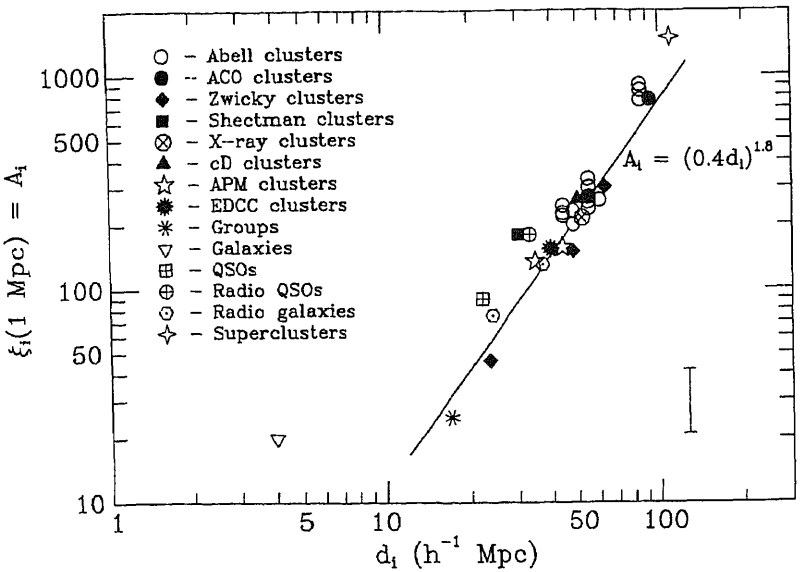


Figure 2. The two-point correlation function amplitude (scaled to 1 Mpc) between objects belonging to a given richness class is plotted against their mean separation d_i. Reproduced, with permission, from [29].

in clustering amplitude with richness is not well understood at present, although some arguments based on the statistics of Gaussian random fields have been presented [2, 3].

1.2 *Velocity fields of galaxies*

The existence of inhomogeneities in the distribution of galaxies results in a peculiar component to their motion over and above the smooth Hubble flow: $\vec{\mathbf{v}}_{\mathrm{pec}} = \mathbf{V} - H\mathbf{r}$ ($\mathbf{V}$ is the total galactic velocity and $H\mathbf{r}$ is the Hubble flow). Estimates of the peculiar velocity depend upon the sample depth, most surveys indicate $\vec{\mathbf{v}}_{\mathrm{pec}} \simeq 350\,\mathrm{km/sec}$, on scales $\sim 50\,\mathrm{Mpc}$. Direct evidence for the motion of our galaxy arises from the dipole anisotropy of the cosmic microwave background $\Delta T/T \simeq 1.2 \times 10^{-3}$, which is usually taken to indicate that the milky way is moving at a velocity $\vec{\mathbf{v}}_{\mathrm{pec}} \simeq 610\,\mathrm{km/sec}$ relative to the MBR. A study of large scale motions of elliptical galaxies made by Lynden–Bell *et al* [4] indicated that a sample of galaxies spanning a volume $\sim (50\,h^{-1})^3$ cubic megaparsec could be participating in a bulk flow directed towards Centaurus. The large amplitude of the flow ($\sim 570 \pm 60\,\mathrm{km/s}$ in the vicinity of the milky way) led some theorists to suggest that a large density fluctuation termed 'the great attractor' (GA) lying in the direction of Centaurus was responsible for this motion. Attempts to 'see' the great attractor, have not yet yielded definite results mainly because the GA region lies in a zone partly obscured by the plane of the milky way. More recent analysis of velocity data obtained from the IRAS satellite indicate that the peculiar velocity of the milky way may be due to the combined gravitational pull of a dozen or so clusters including Virgo, Hydra and Centaurus. If this is indeed the case, then the notion of a GA appears unnecessary.

Since bulk flows in galaxies are directly related to fluctuations in the gravitational potential, considerable effort in the past few years has been paid to reconstruction methods (such as the 'potent' approach) which attempt to reconstruct the density field from a knowledge of peculiar velocities. Such an approach has the advantage that it maps thedensity field due to *all* matter – dark and visible, and is not biased towards luminous matter as are galaxy catalogues. (A detailed discussion of reconstruction methods can be found in [5].)

1.3 *Cosmic microwave background anisotropy*

A major astrophysical discovery of recent years has been the detection of large angle fluctuations ($> 7°$) in the cosmic microwave background (CMB) made by the COBE satellite in 1992. (These fluctuations are thought to have a primordial origin unlike the much larger dipole fluctuation discussed earlier whose presence is attributed to the peculiar motion of our galaxy.) CMB fluctuations are related to fluctuations in matter at the time of the cosmological recombination of hydrogen (which also marks the decoupling of matter and radiation). The small amplitude of these fluctuations ($\Delta T/T \simeq 10^{-5}$) indicates that the universe was fairly homogeneous and isotropic on scales larger than the cosmological particle horizon, at the time of recombination. The spectrum of the anisotropy implies a primordial power spectrum for matter $P(k) = Ak^n$, $n = 1.1 \pm 0.32$ [6], which is in good agreement with predictions made by the inflationary scenario over a decade ago [7]. The COBE results lend further support to the 'standard model' of structure formation whereby the large scale structure in the universe forms due to the gravitational instability of density fluctuations which grow from initially small values $\delta\rho/\rho|_{\text{initial}} \sim \Delta T/T \simeq 10^{-5}$ at the time of recombination, to $\delta\rho/\rho|_{\text{final}} \gg 1$ today.

2. Nonlinear approximations to gravitational clustering

Gravitational instability in a collisionless self-gravitating fluid can be modelled by the coupled Euler–Poisson system of partial differential equations. These equations are written in terms of $\mathbf{x} = \mathbf{r}/a$ (the comoving spatial coordinate), $\mathbf{v} = \dot{\mathbf{r}} - H\mathbf{r} = a\dot{\mathbf{x}}$ (the peculiar velocity field), $\phi(\mathbf{x}, t)$ (the peculiar Newtonian gravitational potential) and $\rho(\mathbf{x}, t)$ (the matter density). Neglecting pressure terms we obtain [8]

The Euler equation

$$\frac{\partial(a\mathbf{v})}{\partial t} + (\mathbf{v} \cdot \nabla_x)\mathbf{v} = -\nabla_x \phi. \tag{5}$$

The term in the right hand side of eq. (5) represents the gravitational force which gives rise to acceleration in the fluid element.

The continuity equation

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_x[(1+\delta)\mathbf{v}] = 0, \tag{6}$$

where $\delta = \rho/\rho_0 - 1$ is the *density contrast* (ρ_0 is the mean background density), and

$$\nabla_x^2\phi = 4\pi G a^2(\rho - \rho_0) = 4\pi G a^2 \rho_0 \delta. \tag{7}$$

The Euler–Poisson system forms a strongly coupled non-linear system of differential equations which is difficult to treat analytically. Several approximation schemes have been suggested which partially remedy this problem, we describe some of them below.

2.1 *Perturbation theory*

During the early stages of gravitational clustering the density contrast is small $\delta \ll 1$ and one can attempt to solve eq. (5)–(7) perturbatively by writing the density contrast as a series

$$\begin{aligned}\delta(\mathbf{x}, t) &= \sum_{n=1}^{\infty} \delta^{(n)}(\mathbf{x}, t) \\ &= \sum_{i=1}^{\infty} D_+^n(t)\Delta^{(n)}(\mathbf{x}).\end{aligned} \tag{8}$$

The first term in the above expansion corresponds to the linear growth law: $\delta^{(1)} = D_+(t)\Delta^{(1)}(\mathbf{x})$. In a spatially flat matter dominated universe $D_+(t) \propto a(t) \propto t^{2/3}$ where $a(t)$ is the scale factor of the universe [9]. (The above expression tells us that left to themselves, density fluctuations would grow by a factor $\simeq 10^3$ between the epoch of decoupling and the present epoch.) Perturbation theory formally breaks down when $\delta \simeq 1$, to follow gravitational instability. Further one has to use more sophisticated non-linear approximations such as the Zel'dovich approximation described below.

2.2 *The Zel'dovich approximation*

In 1970, Zel'dovich proposed a remarkably simple approximation to follow the motion of fluid elements into the strongly non-linear regime. Zel'dovich suggested that instead of perturbing the density field one should instead perturb particle trajectories. The resulting Zel'dovich approximation (ZA) has the simple form [10, 11]:

$$\mathbf{r} = a\mathbf{x}(\mathbf{q}, t) = a(t)[\mathbf{q} + D_+(t)\mathbf{u}(\mathbf{q})], \tag{9}$$

here $\mathbf{q}$ is the initial (Lagrangian) coordinate of a particle, and $\mathbf{x}$ is its final (Eulerian) coordinate. $\mathbf{u}(\mathbf{q})$ is the initial (comoving) velocity field of the particle at $\mathbf{q}$: $\mathbf{u}(\mathbf{q}) = \mathbf{v}/a\dot{D}_+$. Since any initial rotation is strongly damped during expansion we restrict our attention to potential flows so that $\mathbf{u}(\mathbf{q}) = -\nabla\Phi_0(\mathbf{q})$, $\Phi_0(\mathbf{q})$ being the linear velocity potential. It can be shown that Φ_0 is related to the Newtonian gravitational potential ϕ by $\phi = A^{-1}\Phi_0$ where $A = 2/(3H^2a^3)$; A is a constant for a spatially flat matter dominated universe.

A generic feature of the Zel'dovich approximation (for random initial distributions of Φ_0) is the intersection of particle trajectories (shell crossing) which formally leads to singularities in the density field. As Zeldovich pointed out, the motion described by

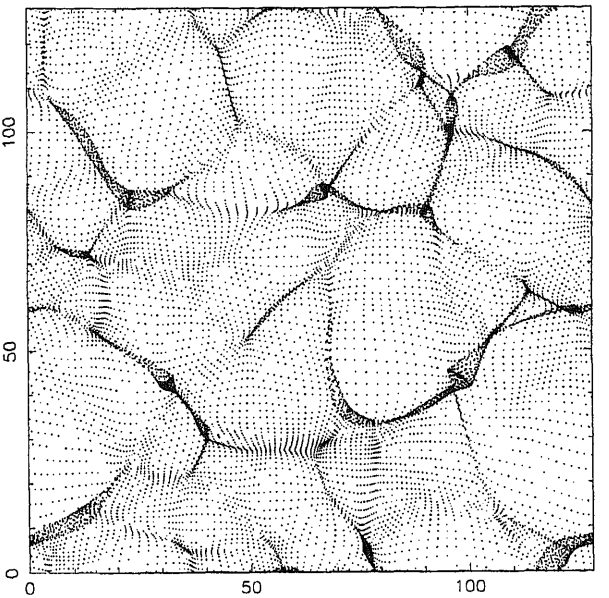


Figure 3. Caustics arising in a two dimensional realisation of the Zel'dovich approximation are shown for random (Gaussian) initial conditions. Reproduced, with permission, from [8].

eq. (9) is in many respects very similar to the equation describing the propagation of light rays in geometrical optics [11]. In optics, caustics – describing alternating bright and dark regions in the intensity of light – form from generic initial conditions (all of us are accustomed to seeing such features at the bottom of a shallow pool of water). Similarly gravitational instability too – as predicted by the Zel'dovich approximation – will result in the formation of caustics in the density field, as shown in figure 3.

The prediction of *planar caustics* follows immediately from eq. (9) if we note that the condition for mass conservation

$$\mathrm{d}M = \rho_0 \mathrm{d}^3\mathbf{q} = \rho(\mathbf{x}, t)\mathrm{d}^3\mathbf{x} \tag{10}$$

leads to the following expression for the density in terms of the Jacobian of the transformation from $\mathbf{q}$ to $\mathbf{x}$:

$$\begin{aligned}\frac{\mathrm{d}V_e}{\mathrm{d}V_l} &= J\left[\frac{\partial \mathbf{x}}{\partial \mathbf{q}}\right] \\ &= \left|\delta_{ij} - D_+(t)\frac{\partial^2 \Phi_0}{\partial q_i \partial q_j}\right| \\ &= [1 - D_+(t)\lambda_1(\mathbf{q})][1 - D_+(t)\lambda_2(\mathbf{q})][1 - D_+(t)\lambda_3(\mathbf{q})],\end{aligned} \tag{11}$$

where $\mathrm{d}V_e \equiv \mathrm{d}^3\mathbf{x}$ and $\mathrm{d}V_l \equiv \mathrm{d}^3\mathbf{q}$, are the Eulerian and Lagrangian volume elements, respectively, and λ_1, λ_2 and λ_3 are the eigenvalues of the deformation tensor $d_{ij} \equiv \partial^2\Phi_0/\partial q_i \partial q_j$.

The resulting expression for the density distribution is

$$\rho(\mathbf{r},t)=\frac{\rho_0}{a^3}[1-D_+(t)\lambda_1(\mathbf{q})]^{-1}[1-D_+(t)\lambda_2(\mathbf{q})]^{-1}[1-D_+(t)\lambda_3(\mathbf{q})]^{-1}. \quad (12)$$

For a Gaussian random field the probability for the largest eigenvalue to be positive is 92%. A volume element satisfying this criterion will collapse along at least one direction, resulting in the formation of two dimensional pancake-type singularities in the density field, which arise when $1-D_+(t)\lambda_1=0$; (the eigenvalues are assumed to be ordered $\lambda_1>\lambda_2>\lambda_3$) see figure 3. Thus Zel'dovich showed that the first singularities in the density field are likely to be planar, a development which led to the 'pancake' model of structure formation pioneered by Zel'dovich and his collaborator's in Moscow in the 1970's and 1980's.

An interesting feature of ZA is that caustics percolate, even though they occupy a negligible fraction of the total volume. For a Gaussian random field a phase has to occupy at least 16% of the total volume in order to percolate. The reason for percolation in ZA is simple, regions which collapse in ZA occupy 92% of the Lagrangian volume, guaranteeing that such regions will percolate in Lagrangian space. Since the Zel'dovich mapping (9) preserves topology, regions which percolate in Lagrangian space continue to do so after they are mapped into Eulerian space (even though they now occupy a much smaller volume). One sees a similar property in the large scale distribution of galaxies, which also percolates at a small value of the 'filling factor'. Thus the Zel'dovich approximation could be on the right track in describing the gravitational clustering process that led to the formation of large scale structure in the universe.

2.3 *The adhesion model*

Tests of the Zel'dovich approximation against the results of N-body simulations show that the approximation works exceedingly well until the formation of the first pancakes. The Zel'dovich approximation attempts to describe particle motion 'inertially' – entirely through the initial conditions. There is therefore no inbuilt feedback mechanism which can modify particle trajectories especially when they pass through high density regions such as caustics. This drawback in ZA results in its breakdown soon after the formation of caustics which are shown in figure 3. If one applies ZA naively to later epochs one will find that the thickness of pancakes grows rapidly as particles sail through them oblivious of the backreaction of strong gravitational fields which they encounter enroute. Contrary to this one finds that the thickness of pancakes soon stabilises in N-body simulations. A successful attempt to incorporate the 'adhesive' effects of gravity on small scales into the Zel'dovich approximation has resulted in the adhesion model of structure formation (AM) [11, 12]. AM allows the study of large scale structure to be carried out in the strongly nonlinear regime after the first pancakes have formed, it reduces to ZA at early times. In AM particles are made to stick together once they enter pancakes, so that the total velocity flow is conserved. To appreciate the formal structure of AM let us recast the Euler equation in terms of the comoving velocity field $\mathbf{u}=(\mathbf{v}/\dot{a}a)=\mathrm{d}\mathbf{x}/\mathrm{d}a$ and the comoving velocity potential $\mathbf{u}=-\nabla\Phi$. As a result equation (5) reduces to

$$\frac{\partial\mathbf{u}}{\partial a}+(\mathbf{u}\nabla_x)\mathbf{u}=-\frac{3}{2a}[A\nabla_x\phi+\mathbf{u}], \quad (13)$$

$$\frac{\partial \Phi}{\partial a} - \frac{1}{2}(\nabla_x \Phi)^2 = \frac{3}{2a}(A\phi - \Phi). \tag{14}$$

In the linear regime of gravitational instability $\Phi = A\phi$, leading to

$$\frac{\partial \Phi}{\partial a} - \frac{1}{2}(\nabla_x \Phi)^2 = 0 \tag{15}$$

which has the analytical solution

$$\Phi(\mathbf{x}, a) = \Phi_0(\mathbf{q}) - \frac{(\mathbf{x} - \mathbf{q})^2}{2a}. \tag{16}$$

Differentiating eq. (16) alternately with respect to $\mathbf{x}$ and $\mathbf{q}$, we get the Zel'dovich approximation

$$\begin{aligned} \mathbf{x} &= \mathbf{q} - a(t)\nabla_q \Phi \\ \mathbf{u} &= -\nabla_x \Phi(\mathbf{x}, a) = -\nabla_q \Phi_0(\mathbf{q}). \end{aligned} \tag{17}$$

We thus see that the Zel'dovich approximation was inspired by the linear relationship $\Phi = A\phi$ which was successfully extended into the non-linear regime. (As a consequence of this relationship the velocity vector remains parallel to the acceleration vector so that particles move along straight line trajectories in ZA.) The adhesion approximation simplifies equation (14) by assuming that the right hand side in that equation can be replaced by a mock viscosity term intended to mimick the 'adhesive' effects of gravity on small scales. The resulting Euler equation becomes [12, 13]

$$\frac{\partial \mathbf{u}}{\partial a} + (\mathbf{u}\nabla)\mathbf{u} = \nu \nabla^2 \mathbf{u}. \tag{18}$$

This equation is the three dimensional generalization of Burgers' equation [14, 15]. If the motion happens to be potential $\mathbf{u} = -\nabla\Phi$, eq. (18) can be solved analytically by means of the Hopf–Cole substitution $\Phi(\mathbf{x}, a) = -2\nu \log U(\mathbf{x}, a)$ which leads to the diffusion equation

$$\frac{\partial U}{\partial a} = \nu \nabla^2 U \tag{19}$$

which can be solved exactly. We therefore get

$$\mathbf{u}(\mathbf{x}, a) = \frac{\int \mathrm{d}^3\mathbf{q}((\mathbf{x} - \mathbf{q})/a) \exp[-S(\mathbf{x}, a; \mathbf{q})/2\nu]}{\int \mathrm{d}^3\mathbf{q} \exp[-S(\mathbf{x}, a; \mathbf{q})/2\nu]}, \tag{20}$$

where

$$S(\mathbf{x}, a; \mathbf{q}) = -\Phi_0(\mathbf{q}) + \frac{(\mathbf{x} - \mathbf{q})^2}{2a(t)}. \tag{21}$$

The trajectory of a particle can be determined from the integral equation [16, 17]

$$\mathbf{x}(\mathbf{q}, a) = \mathbf{q} + \int_0^a \mathrm{d}a' \mathbf{u}[\mathbf{x}(\mathbf{q}, a'), a']. \tag{22}$$

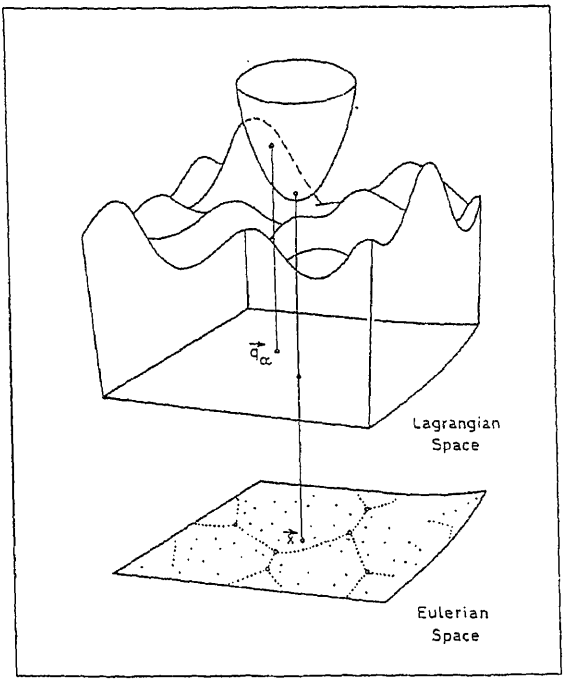


Figure 4. The geometrical construction of the adhesion model. A paraboloid descends onto the linear gravitational potential and defines sites of structure formation in two dimensions. The paraboloid is tangent to the potential in *Lagrangian* space, its apex gives the location of two dimensional caustics (filaments and clumps) in *Eulerian* space. Reproduced, with permission, from [8].

For cosmological applications it is necessary to be able to solve eqs (20), (22) in the limit $\nu \to 0$. This can be done using the method of steepest descents which indicates that the maximum contribution to the integral in eq. (20) will come from those points where $\nabla_q S = 0$. This results in an interesting geometrical interpretation which we describe below (see [18, 19]). We notice that the condition $\nabla_q S = 0$ defines a paraboloid $P = (\mathbf{x} - \mathbf{q})^2/2a(t)$ tangential to the velocity potential Φ_0 at $\mathbf{q}$. The large scale structure of the universe is obtained by descending a paraboloid P onto the (initial) potential Φ_0 (P and Φ_0 are not permitted to intersect). The apex of P defines the final location of particles which are initially located at $\mathbf{q}$ (the point of contact between P and Φ_0). The mapping between the initial position $\mathbf{q}$ of a particle (also called its Lagrangian coordinate), and its final location $\mathbf{x}$ (Eulerian coordinate) is single valued at early times when $a(t)$ (the radius of curvature of P) is small. It is multi-valued at late times when $a(t)$ is large. This means that the velocity flow is *single stream* at early times (i.e. caustics have not yet formed), but *multistream* at late times (after the formation of caustics). It is easy to see that for small a the geometrical approach reduces to the Zel'dovich approximation. However the full power of the adhesion approach is felt at late times when the Zel'dovich approximation is no longer valid. The adhesion model at such times tesselates the universe into pancakes (or sheets), filaments and clumps

complemented by large empty regions – voids (see figure 4). Sheets form at the apex of the paraboloid P when it touches Φ_0 in exacly two points; filaments form when P touches Φ_0 in three points and clumps when the points of contact are exactly four. (More than four contact points are not generic in three dimensions.) According to the adhesion model gravitational instability proceeds in several distinct stages: during the first particles move 'inertially' to form two-dimensional 'pancakes', within pancakes flows become multistream and matter aggregates in regions of higher densities – filaments (which form at the intersections of pancakes). Finally matter flows along filaments to collect in clumps (the intersections of filaments). Thereafter neighboring clumps attract and merge and gravitational clustering proceeds in a hierarchical manner.

The adhesion model permits us to view the formation of large scale structure in the universe in a geometrical manner. The picture of clustering at different epochs is obtained by varying the radius of curvature of the paraboloid which is simply the scale factor of the universe. Although AM does not provide details of fluid motion within caustics it does give a fairly accurate description of structure on large scales as verified by detailed tests against N-body simulations [17, 20].

3. Percolation and the shape of large scale structure

Dynamical methods which describe how structures form must be supplemented by statistical methods describing different aspects of gravitational clustering. The simplest statistic to be applied to galaxy catalogues is the two point correlation function ξ which has been described earlier. Since ξ is the Fourier transform of the power spectrum $P(k) = |\delta_k|^2$ it ignores information contained in the phases of Fourier components δ_k. This is a major limitation since it is well-known that beyond the linear regime of gravitational instability coupling between modes becomes important leading to the build up of phase correlations. (Initially the phases of the Fourier components of the density field are uncorrelated and random.) A consequence of mode coupling is the departure of the galaxy distribution from Gaussianity and the increasing coherence length of structures. (In the universe one sees 'patterns' in the distribution of galaxies on scales of 50–100 Mpc. i.e. on scales on which $\xi(R)$ is effectively zero.) Large scale structure in the universe appears to be concentrated in large objects (such as the *great wall*) occupying a relatively small volume of space (i.e. the filling factor of such objects is very small), see figure 1. These 'superclusters' of galaxies are separated by voids – vast regions of space with little if any matter. Zel'dovich first drew attention to the smallness of the filling factor of the galaxy distribution in the universe and suggested using percolation theory to understand its topological properties [33].

It is well known that a random (Poisson) distribution of points percolates at a filling factor (FF) of ~ 0.319. In other words the fraction of volume occupied by the percolating phase is $\sim 32\%$. On the other hand systems evolving under gravitational instability show a much smaller FF at percolation ~ 0.02–0.07% [30]. This is also what one finds from galaxy surveys leading one to believe that the galaxy distribution is strongly non-Gaussian. In a comprehensive study we have analysed the percolation properties of N-body simulations intending to model the universe at different stages of its evolution.

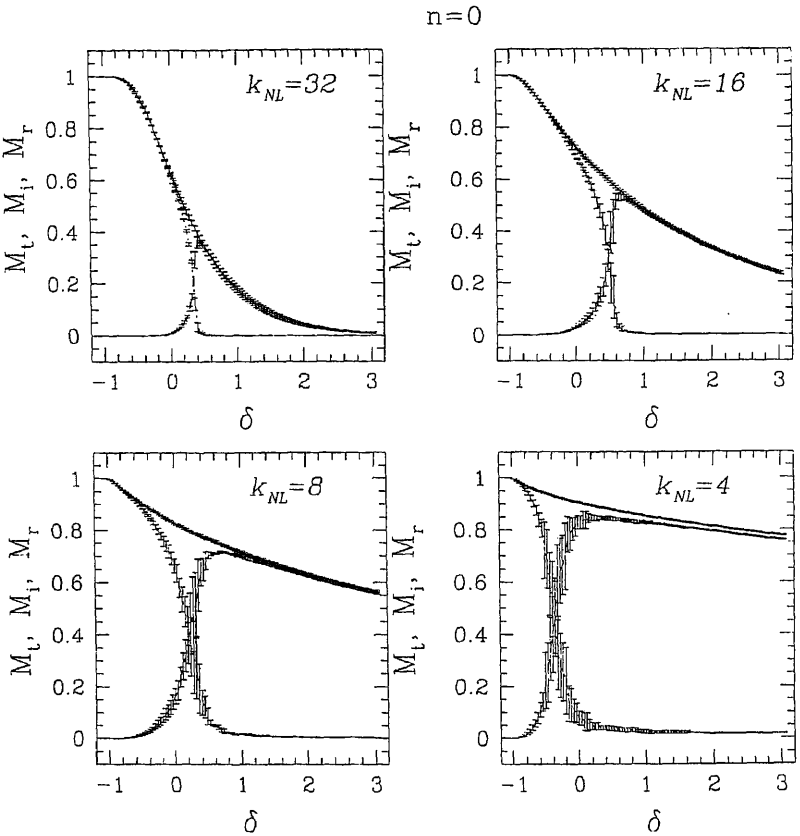


Figure 5. Percolation in $n = 0$ model at different epochs measured by the scale of nonlinearity $k_{\rm NL}^{-1}$ which increases from top to bottom. The filling factor of the largest cluster M_i (dotted line), the filling factor of all but the largest cluster M_r (dashed line) and the total filling factor M_t (solid line) are plotted against the density contrast. Reproduced, with permission, from [21].

In our simulations we first define a density threshold, then identify clusters at this threshold using a nearest neighbor (friends-of-friends) algorithm. Having done this we determine the mass fraction in clusters. The models considered are N-body simulations with power spectrum $P(k) \equiv \langle |\delta_k|^2 \rangle \propto k^n$, for $k < k_N$, and $P(k) = 0$ for $k \geq k_N$, $n = -3, -2, -1, 0$ and $+1$, where k_N is the Nyquist cutoff. $\langle\ \rangle$ denotes the average over an ensemble. The epochs we consider are characterized by the scale of nonlinearity $k_{\rm NL}^{-1}$ at that epoch determined by

$$\sigma \equiv \langle \delta^2 \rangle^{1/2} = D_+(t) \left(4\pi \int_0^{k_{\rm NL}} P(k) k^2 {\rm d}k \right)^{1/2} = 1, \tag{23}$$

$D_+(t)$ being the linear growing-mode of density fluctuations.

Our results are shown in figure 5 for the spectrum $n = 0$, corresponding to an equal distribution of power amongst all modes in the initial conditions (white noise) (a more

comprehensive discussion of our work may be found in [21, 22]). We have plotted: (1) M_i – the mass contained in the largest cluster divided by the total mass in clusters (the largest cluster is formally 'infinite' after percolation sets in), M_i is also called the mass fraction or filling factor (FF) of the largest cluster; (2) M_r – the mass fraction in the rest of the clusters (other than the largest); (3) $M_t = M_i + M_r$ – the total mass fraction in clusters (i.e. the total filling factor). M_i, M_r, M_t are plotted as functions of the density threshold for different expansion epochs in figure 5. We find that M_i decreases with increasing δ, for $\delta > \delta_c \simeq 0.5$, M_i is very small and clusters do not percolate. On the other hand for $\delta < \delta_c$, M_i increases sharply indicating that virtually all the mass (above the threshold) is in the largest (percolating) cluster. Thus $\delta = \delta_c$ defines the percolation threshold which appears to be sensitive both to the spectral index and the amount of evolution the system has undergone. A related issue of considerable importance concerns the *topology* of large scale structure, i.e. whether the distribution of large scale structure is 'meatball-like', 'sponge-like' or some other. High values of the filling factor at percolation (relative to a Gaussian distribution) determine a 'meatball-like' topology, whereas low values indicate a 'bubble-like' or a 'sponge-like' topology [24].

We should note that $M_r \gg M_i$ for $\delta \gg \delta_c$ indicating that most of the mass is in small unconnected clusters at high densities. As we decrease the threshold δ we find that M_r peaks, then rapidly decreases. The fall in M_r is compensated by the increase in M_i, caused by the fact that clusters, as they grow in size, tend to connect with each other to form superclusters which eventually percolate. (The error bars in figure 5 indicate the rms dispersion in four different N-body realisations of the same spectrum.)

The smallness of the filling factor in our simulations indicates that clustered matter is likely to be distributed in either sheets or filaments (since both sheets and filaments occupy less space than spheres). To check whether this is true we employed a statistic sensitive to shape originally suggested by Babul and Starkman [23]. This 'shape statistic' is constructed by determining the moment of inertia (deformation) tensor of matter distributed within a radius R of a fiducial point. The tensor is then diagonalized and a triplet of shape functions S_i introduced by the relations

$$S_1 \equiv \sin\left[\frac{\pi}{2}(1-\mu)^p\right], \tag{24}$$

$$S_2 \equiv \sin\left[\frac{\pi}{2}a(\mu,\nu)\right], \tag{25}$$

$$S_3 \equiv \sin\left[\frac{\pi\nu}{2}\right], \tag{26}$$

where $\mu \equiv \sqrt{\lambda_2/\lambda_1}$, $\mu \equiv \sqrt{\lambda_3/\lambda_1}$. (The three eigenvalues of the deformation tensor λ_1, λ_2, λ_3 are arranged in decreasing order, so that $\lambda_1 > \lambda_2 > \lambda_3$.) The function $a(\mu,\nu)$ is implicitly defined by the equation

$$\frac{\mu^2}{a^2} - \frac{\nu^2}{a^2(1-\alpha a^{1/3}+\beta a^{2/3})} = 1. \tag{27}$$

The parameters p, α and β are chosen to normalize S_n (for details see [23, 8, 21, 22]).

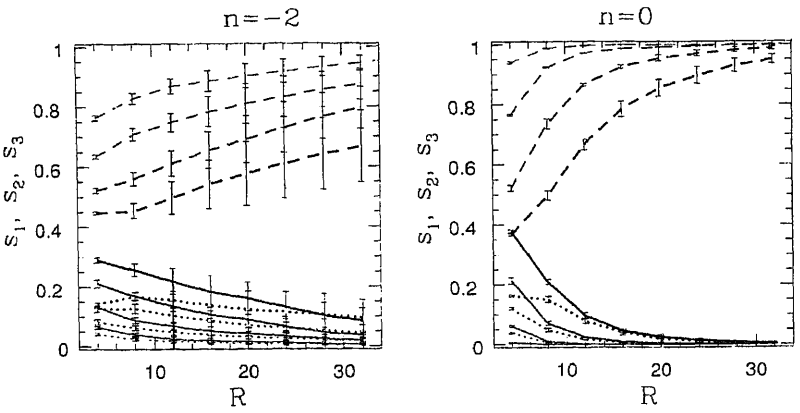


Figure 6. Structure functions averaged over randomly chosen high density regions are plotted as functions of the window radius R (measured in units of the grid size). Solid lines plot the average linearity $\bar{S}_1$, dotted lines the average planarity $\bar{S}_2$, and dashed lines the average sphericity $\bar{S}_3$. Heavier lines correspond to later epochs. The y-axis extends from 0 to 1. Reproduced, with permission, from [22].

Defined in this manner S_i can be evaluated for a distribution of points (as originally suggested in [23]) or for a density field (which is the approach we employ). In the latter case it is sufficient to evaluate $S_i(R)$ within a radius R of fiducial points lying above a given density threshold δ_i. The location of the fiducial points is chosen randomly (the sole criterion being that they lie in regions where $\delta > \delta_i$). One can define the average shape functions $\bar{S}_i = 1/N \sum_{i=1}^{N} S_i$ defined for the ensemble as a whole, and study its properties. S_i can be viewed as a vector, defined in the first octant: $S_i = (1,0,0)$ describes a one dimensional filament, $S_i = (0,1,0)$ – a two dimensional 'panacake', and $S_i = (0,0,1)$ – a three dimensional sphere. In figure 6 we plot $\bar{S}_i$ as functions of both the radius R and the time (characterized by the scale of non-linearity $k_{\rm NL}^{-1}$). We find that as the N-body simulation evolves structures within it proceed to become more and more filamentary. This is true for all spectra we consider and appears to be an important generic property of systems evolving under gravitational instability [31,32].

4. Conclusions

Both dynamical and statistical aspects are crucial to a deep understanding of large scale structure in the universe. Whereas statistical indicators such as the two point correlation function are useful in measuring clustering, they must be complemented by topological descriptors in order to understand the global properties of large scale structure such as its connectedness and topology. We show that gravitationally clustered systems in general percolate at lower filling factor's than a Gaussian random field, this property indicates that the large scale structure in such systems is closer to being 'sponge-like' or 'bubble-like' than 'meatball-like'. We also argue that whereas the first gravitationally bound objects are likely to be pancakes, gravitational clustering thereafter leads to the increasing filamentarity of structures. Whether the degree of filamentarity seen in

simulations matches that in redshift surveys of galaxies (such as IRAS), is presently being studied [27].

Acknowledgement

The author is grateful to B S Sathyaprakash and Sergei Shandarin for stimulating discussions and many fruitful collaborations.

References

[1] M Davis and P J E Peebles, *Astrophys. J.* **267**, 465 (1983)
[2] N Kaiser, *Astrophys. J.* **284**, L9 (1984)
[3] J M Bardeen, J R Bond, N Kaiser and A S Szalay, *Astrophys. J.* **304**, 15 (1986)
[4] D Lynden-Bell, S M Faber, D Burstein, R L Davies, A Dressler, R J Terlevich and G Wegner, *Astrophys. J.* **326**, 19 (1988)
[5] A Dekel, *Ann. Rev. Astron. Astrophys.* **32**, 371 (1994)
[6] M White, D Scott and J Silk, *Ann. Rev. Astron. Astrophys.* **32**, 319 (1994)
[7] A R Liddle and D H Lyth, *Phys. Rep.* **231**, 1 (1993)
[8] V Sahni and P Coles, *Phys. Rep.* **262**, 1 (1995)
[9] In a static universe the growth of δ is much faster $D_+(t) \propto \exp\sqrt{4\pi G\rho t}$
[10] Ya B Zel'dovich, *Astron. Astrophys.* **4**, 84 (1970)
[11] S F Shandarin and Ya B Zel'dovich, *Rev. Mod. Phys.* **61**, 185 (1989)
[12] S N Gurbatov, A I Saichev and S F Shandarin, *Mon. Not. R. Astron. Soc.* **236**, 385 (1989)
[13] S N Gurbatov, A I Saichev and S F Shandarin, *Sov. Phys. Dokl.* **30**, 921 (1985)
[14] J M Burgers, *The non-linear diffusion equation* (Reidel, Dordrecht, 1974)
[15] G B Whitham, *Linear and non-linear waves* (Wiley, New York, 1974)
[16] A Nusser and A Dekel, *Astrophys. J.* **362**, 14 (1990)
[17] D H Weinberg and J E Gunn, *Mon. Not. R. Astron. Soc.* **247**, 260 (1990)
[18] L A Kofman, D Yu Pogosyan, S F Shandarin and A L Melott, *Astrophys. J.* **393**, 437 (1992)
[19] V Sahni, B S Sathyaprakash and S F Shandarin, *Astrophys. J.* **431**, 20 (1994)
[20] B S Sathyaprakash, V Sahni, D Munshi, D Pogosyan and A L Melott, *Mon. Not. R. Astron. Soc.* **275**, 463 (1995)
[21] B S Sathyaprakash, V Sahni and S F Shandarin, *Mon. Not. R. Astron Soc.* (submitted)
[22] B S Sathyaprakash, V Sahni, S F Shandarin, *Astrophy. J. Lett.* **462**, L5 (1996)
[23] A Babul and G D Starkman, *Astrophys. J.* **401**, 28 (1992)
[24] In a meatball topology underdense regions percolate easily whereas overdense do not. In a bubble-topology the underdense phase is enclosed by overdense 'bubbles', as a result the overdense phase percolates whereas the underdense one does not. In a sponge topology both phases are symmetric and both can percolate with equal ease; (see [8,25] for a detailed discussion of these issues). Other statistical indicators used to investigate the global properties of large scale structure include the 'genus characteristic' and the minimal spanning tree [8,26].
[25] C Yess and S F Shandarin, *Astrophys. J.* **465**, 2 (1996)
[26] A L Melott, *Phys. Rep.* **193**, 1 (1990)
[27] B S Sathyaprakash, V Sahni, S F Shandarin and K Fisher, in preparation
[28] L N da Costa, in *Cosmic velocity fields*, Proceedings of the ninth IAP Astrophysics meeting edited by F R Bouchet and M Lachieze-Rey (Editions Frontieres 1994) p. 475
[29] N A Bahcall and M J West, *Astrophys. J.* **392**, 419 (1992)
[30] A A Klypin and S F Shandarin, *Astrophys. J.* **413**, 48 (1993)
[31] R C Pearson and P Coles, *Mon. Not. R. Astron. Soc.* **272**, 231 (1995)
[32] B S Sathyaprakash, V Sahni, S F Shandarin and K Fisher, in preparation
[33] Ya B Zel'dovich, *Sov. Astron. Lett.* **8**, 102 (1982)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 147–154

Rotation curves of galaxies: Missing mass or missing physics

V KRISHAN
Indian Institute of Astrophysics, Bangalore 560 034, India

Abstract. The rotation curves of galaxies are modelled using very special properties of an hydrodynamically turbulent fluid possessing helicity fluctuations. The development of correlations among these fluctuations leads to the formation of organized structures characterized by a new flat branch of the spatial energy spectrum in addition to the well known Kolmogorov spectrum. It is proposed that the flat nature of the rotation curves of galaxies may be a result of the energy cascading processes occuring in turbulent galactic atmospheres. Thus, in this model, there is no need of invoking dark matter to account for the flat rotation curves of galaxies.

Keywords. Rotation curves; galaxies; turbulence; inverse cascade.

PACS Nos 95.30; 98.30; 98.52

1. Introduction

We observe coherent structures and correlated motions delineating well defined patterns on all spatial and temporal scales. Some of the examples are: a convection cell, a tornado, a cyclone, convective flows on stellar surfaces, the Red spot on Jupiter, spiral patterns in galaxies, clusters of galaxies and perhaps – ourselves. The belief is that at the roots of this diversity lie the unique and transcendental properties of non-equilibrium systems. Such systems exhibit organization of matter and motion through the formation of dissipative structures. A stable system near equilibrium, when disturbed bounces back to the initial state. But a system driven sufficiently far from equilibrium may become unstable, reach a bifurcation point after which it attains a completely new identity, a dissipative structure, named so because it needs more energy to sustain itself than the initial configuration. Associated with these bifurcations is the breaking of symmetry, i.e. one starts with symmetric equations and ends up with asymmetric solutions, one of which the system chooses to obey. Sea shells often show a preferential chirality! That means, we need representations of dynamics, which are not invariant with respect to time inversion. Such representations have been found for highly unstable systems [1]. Thus, structure formation takes place in an intrinsically random, irreversible, unsymmetric non-equilibrium systems. The essential conditions for the formation of dissipative structures, thus, are: a macroscopic system far from equilibrium, with correlated fluctuations that can maintain themselves by deriving from their environment. Far from equilibrium conditions exist everywhere in the universe, the big bang being the ultimate example. Many structures like clusters of galaxies are suggested to form and exist in non-equilibrium circumstances [2]. Attempts to fathom them in equilibrium have lead us to darkness; to an unending search and

an unquenchable need for dark matter! Here, we show that the flat rotation curves of galaxies can be successfully modeled if the recently discovered properties of a turbulent medium are included in addition to the standard gravitational effects [3]. Fluid turbulence is a complex subject. Even a genius like Heisenberg seems to have said that he hoped, before he died, someone would explain quantum mechanics to him, but after he died, he hoped, God, would explain turbulence to him [4]. Simply put, turbulence is a random state of fluid motions on many different spatial and temporal scales exchanging energy among themselves. The spatial and temporal scales are constrained by boundaries, buoyancy and dissipation. The problem of turbulence is addressed in two ways: (i) the Kolmogorov approach in which the statistically stationary states are studied using dimensional arguments and (ii) the Navier–Stokes way, in which, one hopes that the solutions of the Navier–Stokes equations would comply with the predictions of the Kolmogorov approach [5, 6]. Here, we describe the Kolmogorov approach.

2. The Kolmogorov approach

It is well known that the energy cascades from large spatial scales to small spatial scales in a homogeneous and isotropic three dimensional fluid turbulence and the energy spectrum is given as $E(K)\alpha K^{-5/3}$. But, what if the assumptions of homogeneity and isotropy are dropped? In two-dimensional turbulence, the energy

$$E = \int \frac{1}{2}\rho V^2 \mathrm{d}^3 r \tag{1}$$

and the enstrophy

$$U = \int (\nabla \times \mathbf{V})^2 \mathrm{d}^3 r \tag{2}$$

are the two quadratic inviscid invariants for vanishing normal component of velocity $\mathbf{V}$ i.e. for rigid or periodic boundaries. The Kolmogorovic arguments give energy spectrum in the inertial range, corresponding to the energy invariant as

$$E(K) \propto K^{-5/3} \tag{3}$$

and corresponding to the enstrophy invariant as

$$E(K) \propto K^{-3} \tag{4}$$

It has been verified by several means that the energy cascades from small scales to large scales according to the spectrum given by equation (4) and from large scales to small scales according to the spectrum given by equation (3). Thus a 2-D system admits an inverse cascade of energy. This is also true for a quasi 2-D system i.e. the one with the vertical component of velocity and spatial scale much smaller than their horizontal components [7, 8]. But the question is: is inverse cascade of energy possible in a 3-D system? Since, the real world is 3-dimensional! The recent developments in 3-D turbulence point to the existence of inverse cascade under well defined conditions. Using

the standard fluid equations

$$\frac{\mathrm{d}\mathbf{V}}{\mathrm{d}t} = -\frac{1}{\rho}\nabla p + \mathbf{F}, \tag{5}$$
$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0,$$
$$p = p(\rho), \quad \mathbf{F} = -\nabla \varphi,$$

it can be shown that the system supports another invariant called helicity H defined as

$$H(t) = \int \mathbf{V} \cdot (\nabla \times \mathbf{V}) \mathrm{d}^3 r \tag{6}$$

under vanishing normal component of the vorticity $\vec{\omega} = \nabla \times \mathbf{V}$. Turbulence with $H \neq 0$ lacks reflectional symmetry. For a reflectionally symmetric system, $H = 0$, but it may happen that the higher moments of helicity distribution are non-zero and affect the statistics of the flow [9]. The moments of helicity distribution are defined as

$$H_n = \lim_{V_L \to \infty} \frac{1}{V_L} \Sigma_i [h^{(i)}]^n,$$

where

$$h^i = \int_{V_i} (\mathbf{V} \cdot \nabla \times \mathbf{V}) \mathrm{d}^3 r_i \tag{7}$$

and V_i is a small volume element of the total volume V_L. For a random distribution, $H_1 = 0$ but all even moments are finite and in particular $H_2 \neq 0$ i.e. the fluctuations about the mean have constant variance. Thus, a new invariant I is defined as

$$I = \int \langle (\mathbf{V}_1 \cdot \nabla \times \mathbf{V}_1)(\mathbf{V}_2 \cdot \nabla \times \mathbf{V}_2) \rangle \mathrm{d}^3 r \simeq A \int E^2(K) \mathrm{d}K \tag{8}$$

which describes helicity–helicity correlations and can be expressed in terms of the energy spectrum $E(K)$ for a quasi-normal distribution of helicities. Here $E = \int E(K) \mathrm{d}K$ is the total energy density.

In the inertial range for the energy invariant we have

$$(KV_K)V_K^2 = \epsilon = V_0^2 \tau,$$

where K = wavenumber, V_0 = the initial rms velocity on small scales, τ = the duration for which this energy is available, V_K = velocity in Fourier space and ϵ = average energy transfer rate per unit mass (ergs $\mathrm{g}^{-1}\,\mathrm{s}^{-1}$). This, combined with $KE(K) = V_K^2$ yields the well-known Kolmogorov spectrum

$$E(K) = \epsilon^{2/3} K^{-5/3}. \tag{9}$$

It would be appropriate to comment on ϵ here. Kolmogorov (1941) conjectured that in a quasi-steady state there should be a stationary flow of energy in the K space from the source to the sink. Thus the energy transfer rate per unit mass should be a constant and be equal to the dissipation rate at the sink. Although numerous experiments have confirmed

that ϵ is a strongly fluctuating quantity, surprisingly there is no experimental evidence indicating a deviation from the Kolmogorov spectrum [10].

The value of ϵ for the galaxy has been estimated to be of the order of 8×10^{-3}ergs g^{-1} s^{-1} by considering the various sources (such as supernovae, stellar winds, etc) which contribute to the turbulence energetics. In the same vein τ is calculated to be 3×10^7 yr [11].

From equation (5), we find total energy $E = \int E(K)\mathrm{d}K$ or

$$E(l) = \epsilon^{2/3} l^{2/3} \quad (\text{for } K \simeq 1/l).$$

The corresponding velocity field may be described as

$$V(l) = (l_z \epsilon)^{1/3} (l/l_z)^{1/3} \tag{10}$$

for some normalizing length, l_z. Similarly, in the inertial range for the I-invariant, we have

$$(KV_K)[KE^2(K)] = \epsilon' = I_0/\tau, \tag{11}$$

where ϵ' = the average mean square helicity density exchange rate between the scales. Combining this with

$$KE(K) = V_K^2 \tag{12}$$

gives

$$E(K) = (I_0/\tau)^{2/5} K^{-1}.$$

Or in real space,

$$E(l) = (I_0/\tau)^{2/5} \ln(l/l_z) \quad \text{for } l > l_z. \tag{13}$$

Here, the normalizing length l_z can be used to make the transition from one inertial law (eq. (10)) to the other (eq. (13)). The velocity field in this range is described as

$$V(l) = (\epsilon^2 l_z \tau)^{1/5} [\sqrt{\ln(l/l_z)}], \tag{14}$$

where

$$I_0 = V_0^4 l_z$$

which follows from eqs (11) and (12).

3. Modeling of rotation curves

The issue of the flat rotation curves of galaxies and the need for dark matter is described very precisely in figure 1 [12]. The flat nature of the orbital motion in galaxies is accounted through the relationship $(mV^2)/2 = GMm/R$ by assuming that $M \propto R$ and therefore velocity V = constant. Since the mass $M \propto R$ has no luminosity associated with it, it is known as dark matter. We present an alternative explanation of a flat rotation curve, by resorting to some properties of turbulence described in the previous section.

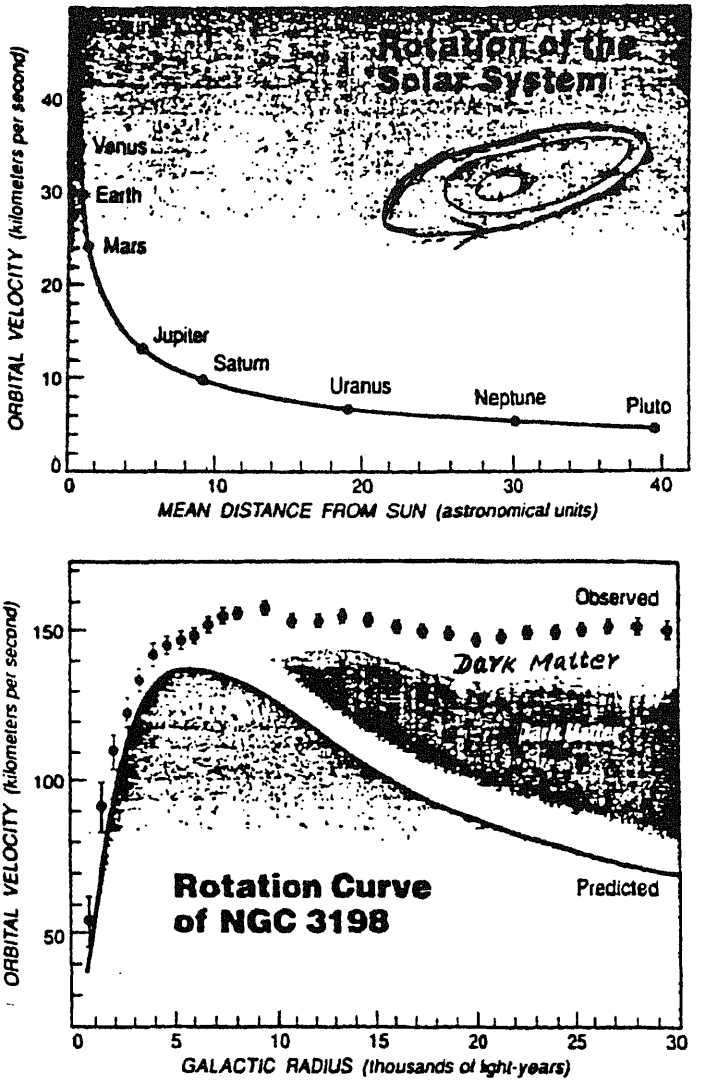


Figure 1. The case for dark matter in spiral galaxies. Top: The orbital velocities of the planets (dots) decrease with distance from the Sun exactly as predicted by Newtonian gravitation (line), assuming a system dominated by one solar mass at its center. Bottom: The cosmos is not as well behaved on galactic scales. Here a graph of orbital velocity versus radius has been computed for NGC 3198, a spiral galaxy in Ursa Major, assuming that the distribution of light serves as a good indicator of the distribution of mass. The failure of the observed velocities (dots) to match the predicted ones is striking and points to an unseen component of dark matter in the galaxy. Courtesy the author.

The complete energy spectrum in a helically turbulent medium derived in Krishan [8] and Krishan and Sivaram [13] is reproduced here in figure 2. We have modeled the rotation curves of 21 galaxies observed by Amram *et al* [14], using the Kolmogorov

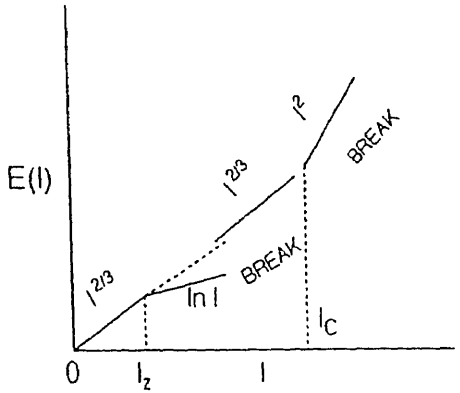


Figure 2. Turbulent energy spectrum. l_z, normalizing length; l_c, break due to Coriolis force [15].

branch [$V(l) \propto l^{1/3}$] and the flat branch [$V(l) \propto \sqrt{\ln l}$]. We propose a law of velocities which is of the type

$$V(l) = Al + Bl^{1/3} \tag{15}$$

in the inner, i.e., $l \leq l_z$, and

$$V(l) = Cl^{-1/2} + D\sqrt{\ln(l/l_z)} \tag{16}$$

in the outer regions, i.e., $l \geq l_z$ of a galaxy, where A, B, C and D are the coefficients to be determined from the fits, with the observed velocity fields.

The first terms on the right-hand side of eqs (15) and (16) correspond to rigid rotation and gravity, respectively; therefore

$$A = \omega,$$

the angular velocity of a galaxy, and

$$C = \sqrt{GM},$$

where G is the universal gravitational constant, refers to the mass of a galaxy. The second terms on the right-hand side of eqs (15) and (16) are due to the turbulence cascading so that

$$B = \epsilon^{1/3}$$

and

$$D = (\epsilon^2 l_z \tau)^{1/5}.$$

By a judicious choice of l_z we can estimate $V_0, \tau, \epsilon, \omega$, and mass M of a galaxy. Some of the modeled rotation curves are shown in figure 3.

Our model gives typical values of the various quantities as

$$V_0 \approx 100\,\mathrm{km\,s^{-1}},$$
$$\tau \approx 10^7\,\mathrm{yr},$$

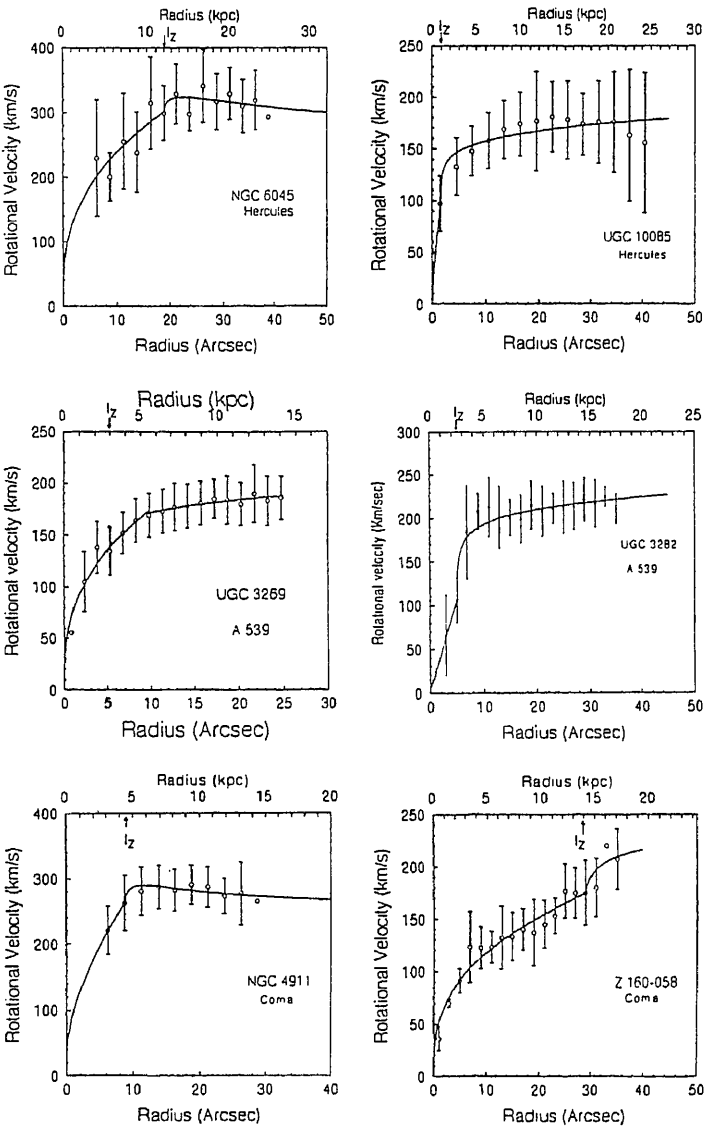


Figure 3. Rotation curves of the galaxies.

$$\epsilon \approx 10^{-2}\,\mathrm{ergs\,g^{-1}\,s^{-1}},$$

$$\omega \approx 10^{-16}\,\mathrm{s^{-1}},$$

$$\text{Mass} \approx 10^{10}\,\mathrm{M_{\odot}}.$$

One must note that we did not have to choose any abnormal values of l_z for obtaining the best fits and it lies in the range 2–10 kpc. This tells us that on scales smaller than l_z, the turbulence is isotropic and on the scales equal to and large than l_z the turbulence becomes more and more anisotropic facilitating the inverse cascade of energy.

4. Conclusions

The velocity–radius relation has been derived using Kolmogorov arguments. We find that the matter at a large radius exhibits a balance of hydrodynamic forces, i.e., dynamical pressure, and Reynolds stress produced by the forced small-scale flow without the necessity of invoking a gravitational force, generated out of a mass distribution of the type $M \propto R$. In other words, our system is hydrodynamically bound.

We also find that ϵ and τ values for each of the galaxies obtained are almost of the same order as that quoted for our galaxy [11]. Therefore, it appears possible to model the observed rotation curves of galaxies by suitably combining the effects of rigid rotation, gravity, and turbulence. The validity of the 'turbulence model' can be further substantiated by confronting it with the observations of the velocity fields on the larger scales-like clusters and superclusters. It is intriguing that the energy spectrum (figure 1) at the largest scales i.e. $E(L) \propto L^2$ or $V(L) \propto L$ resembles the Hubble flow. Is the Hubble flow, a result of an inverse cascade of energy in a helically turbulent universe?

References

[1] I Prigogine, in *The World of Physics* edited by J H Weaver, vol. 2, p. 689 (1987)
[2] G L Verschuur, *Comment Ap.* **15**, 189 (1991)
[3] R D Prabhu and V Krishan, *Ap. J.* **428**, 483 (1994)
[4] F H Shu, in *The Physics of Astrophysics*, 2 (1991)
[5] V Krishan, *Indian J. Phys.* **B66**, 569 (1992)
[6] V Krishan, *MNRAS* **264**, 257 (1993)
[7] A Hasegawa, *Adv. Phys.* **34**, 1 (1985)
[8] V Krishan, in *Solar Interior and Atmosphere* edited by A N Cox, W C Livingston and M S Matthews (The Univ. of Arizona Press, Tuscon, 1991a) p. 1029
[9] H K Moffat and A Tsinober, *Ann. Rev. Fluid Mech.* **24**, 281 (1992)
[10] E Levich, *Phys. Rep.* **151**, 129 (1987)
E Levich, and E Tzvetkov, *Phys. Rep.* **128**, 1 (1985)
[11] A A Ruzmaikin, A M Shukurov and D D Sokoloff, *Magnetic fields of galaxies* (Dordrecht: Kluwer, 1988), p. 150
[12] The author regrets her inability to find the reference of the published paper from which figure 2 has been reproduced here
[13] V Krishan and C Sivaram, *MNRAS* **250**, 157 (1991)
[14] P Amram, E Le Coarer, M Marcelin, C Balkowski, W T Sullivan III and V Cayatte, *Astronomy & Astrophys. Suppl. Ser.* **94**, 175 (1992)
[15] V Krishan, *MNRAS* **250**, 50 (1991b)
[16] A N Kolmogorov, *Dokl. Akad. Nauk. SSSR* **30**, 299 (1941)
[17] E Levich and E Tzvetkov, *Phys. Rep.* **128**, 1 (1985)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 1,
July 1997
pp. 155–160

Deposition and evaporation of k-mers: Dynamics of a many-state system

MUSTANSIR BARMA
Theoretical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

Abstract. When the dynamics of a system partitions the phase space of configurations into very many disjoint sectors, we are faced with an assignment problem: Given a configuration, how can we tell which sector it belongs to? We study this problem in connection with the dynamics of deposition and evaporation of k particles at a time, from a lattice substrate. For $k \geq 3$, the system shows complex behaviour: (a) The number of disjoint sectors in phase space grows exponentially with the size. (b) The asymptotic time dependence of the autocorrelation function shows slow decays, with power laws which depend on the sector. Both (a) and (b) are explained in terms of a nonlocal construct known as the irreducible string (IS), formed from a particle configuration by applying a deletion algorithm. The IS provides a label for sectors; the multiplicity of possible IS's accounts for (a), and let us determine sector numbers and sizes. The elements of the IS are conserved; thus their motion is responsible for the slow modes of the system, and accounts for (b) as well.

Keywords. Interacting stochastic systems; broken ergodicity; conservation laws.

PACS Nos 02.50; 05.40; 68.45; 82.20

1. Many-state systems

It is well known that the number of microscopic configurations in a large statistical system grows exponentially with the volume. In a lattice system with N sites, the number of configurations $\Omega \sim \exp(\alpha N)$.

If the system is 'simple', it exists in a single macroscopic phase or state, in normal circumstances. Only at exceptional points in the phase diagram which require fine tuning of external parameters like the temperature and pressure, can it exist in one or two or a few more possible phases. The dynamical manifestation of this is that starting with any of the Ω possible configurations, the system will evolve to one or the other of the few steady states, each corresponding to a macroscopic phase.

By contrast, 'complex' systems have a large multitude of equivalent, or almost equivalent phases. The number of macroscopic states grows exponentially $\sim \exp(\gamma N)$. Physical examples of such systems are glasses, spin glasses and rubber. The phase space splits up into an exponentially large number of sectors, each of which consists of configurations which can be reached from each other by the dynamics. Inter-sector transitions do not occur, either because there exist strict conservation laws which forbid transitions, or because there are barriers between sectors, so that the time to go from one

sector to another becomes infinite in the thermodynamic limit. Thus if we were to watch the evolution of the system starting from a configuration C_{initial}, we would find the system evolving into the steady state of the relevant sector. However, since this steady state is only one of a very large number of possibilities, we are faced with an assignment problem: Without explicitly evolving the system in time, how can one tell which of the many steady states will be reached from a given C_{initial}? How is one to label steady states?

In this paper, we discuss a model of k-mers, stochastically depositing on, and evaporating from, a one-dimensional lattice [1,2]. For $k \geq 3$, this system has an exponentially large number of sectors and steady states, and is complex in the sense discussed above. We will show that for this system it is possible to find a sector-labelling scheme which relies on a novel sort of construct called the irreducible string [3]. This construction helps solve the assignment problem posed above. Moreover, we find that there is a nontrivial sector-dependence to the time-dependence of correlation functions: asymptotic decay laws show strong variations from one sector to another [4]. Once again, the irreducible string allows us to understand this sector-wise variation, in terms of the diffusive movement of its conserved but mobile elements.

2. Model of k-mer deposition and evaporation

Our model is a generalization of random sequential adsorption (RSA), a process considered by Flory [5]. In RSA, extended objects (made of k particles each) drop randomly onto a lattice substrate, provided that k adjacent sites on the lattice are empty. Of interest is the filling fraction as a function of time, and spatial correlation functions [6]. Over the past few years several extensions of RSA have been considered [7], including the effect of inter-particle interactions and monomer detachment.

We generalized the process by considering the effect of evaporation, in which k particles at a time may detach from adjacent sites of the substrate, and leave [1,2]. An important aspect of the process is that it could be *any* k adjacent particles which leave — not necessarily the same k particles that fell in together. This property of reconstitution of k-mers is a crucial feature of the dynamics. Thus, in a small time dt, k adjacent empty sites become occupied simultaneously with probability ϵdt (deposition of a k-mer), while k adjacent occupied sites become empty with probability $\epsilon' dt$ (evaporation of a k-mer). Let us associate a pseudospin $S_i = \pm 1$ with site i, the value $+1$ corresponding to an occupied site and -1 corresponding to an empty site. Deposition corresponds to $|\downarrow\downarrow\downarrow\rangle \rightarrow |\uparrow\uparrow\uparrow\rangle$, while evaporation corresponds to the reverse transition.

3. Dimers

The case $k = 2$ corresponds to the deposition and evaporation of dimers. The action of the stochastic dynamics may be represented by an operator which acts in the 2^N-dimensional space of configurations. The spin notation leads to an unexpected identification: if $\epsilon = \epsilon'$, this operator can be mapped to the Hamiltonian of the spin-1/2 Heisenberg model, by making a spin rotation of one sublattice. (Here we have assumed that the lattice is bipartite, i.e., decomposable into two interpenetrating sublattices.) Exploiting the fact that

the total spin operator commutes with the Heisenberg Hamiltonian as a consequence of spin rotational invariance (quite a non-obvious symmetry for the original stochastic process), it is possible to calculate in closed form the steady state autocorrelation function [1,2]

$$C_i(t) = \langle S_i(0)S_i(t)\rangle - \langle S_i\rangle^2.$$

At long times, $C_i(t)$ varies as $t^{-d/2}$ where d is the dimension of the system. If $\epsilon \neq \epsilon'$, the corresponding spin operator has additional terms [2], corresponding to a staggered magnetic field and Dzyaloshniskii–Moriya cross-product interactions [8] of alternating sign. In this case, $C_i(t)$ cannot be evaluated in closed form, but can be shown to decay as $t^{-d/2}$ at large times. This answer is suggestive of a diffusion process at work, and this in turn would require a global conservation law, for all ϵ, ϵ'. In fact it is not hard to see that this is the conservation of $M_A - M_B$, where A and B label sublattices, and $M_A(M_B)$ is the total z-component of magnetization on sublattice $A(B)$. This is because the elementary processes of deposition and evaporation of two adjacent particles evidently increments or decrements the occupation of both sublattices equally.

4. Trimers: Phase-space decomposition

Now consider the case of trimers, $k = 3$ on a one-dimensional lattice of length L. (This case is typical of all higher k.) A major difference from the case of dimers is that the trimer deposition-evaporation model is a many-state system, with the number of sectors in phase space growing exponentially with the size of the system: as we will see below, the full phase space of 2^L configurations is partitioned into $\sim \mu^L$ sectors, where $\mu = 1.618...$ is the golden mean. A second important property of trimer dynamics is that the long-time decay of the steady state autocorrelation function $C_i(t)$ depends on the sector. This aspect is discussed in § 5.

4.1 *The irreducible string*

Both the existence of this large infinity of dynamically disconnected sectors and also the variety of asymptotic decay laws, from sector to sector, can be understood in terms of a novel, nonlocal construct, the irreducible string (IS). The IS is defined as follows [3]. Consider a spin configuration $\mathcal{C} \equiv \{S_1, S_2, \ldots, S_L\}$ on a one-dimensional lattice with free boundary conditions. Starting (say) from the left, delete triplets of successive parallel spins, reducing the size of the spin-string by three by each such reduction. After completing a pass through the full string, re-start from the left and repeat the procedure, again and again. This 'deletion do loop' comes to a halt only when further reduction is impossible. The resulting irreducible string evidently has no parallel triplets of spins.

The irreducible string is a constant of the motion and uniquely labels each sector of phase space. The elementary step of deposition or evaporation changes three successive $\uparrow$ spins to three successive $\downarrow$ spins, or vice versa; evidently, the IS is invariant under this. Also, it is possible to prove that if two configurations have the same IS, they are necessarily reachable from each other by deposition-evaporation dynamics [9].

4.2 *Sector decomposition*

Given this labelling scheme for sectors in phase space, it is possible to prove the following facts about the number of sectors and their sizes [3, 9]:

(i) There are quite a few completely jammed configurations (1×1 sectors), for which the length of the IS is the length L of the lattice. The number of such sectors is $2F_L$, where F_L is the Lth Fibbonacci number, which approaches μ^L as $L \to \infty$.

(ii) The total number of sectors N_L obeys the recursion relation $N_L = N_{L-3} + 2F_L$, as each sector is either completely irreducible, or it is not. Consequently, $N_L \approx (1+\mu)\mu^L$.

(iii) The size $D(L, I)$ of a sector is the number of configurations in it. Here I is the IS which labels it. It turns out that D depends on I only through its length $l(I)$. An explicit formula can be obtained for the generating function $g(I) \equiv \sum_{L=l}^{\infty} x^L D(L, I)$. It is then possible to find the asymptotic growth law for $D(I, L)$ for given I.

(v) The set of all fully reducible configurations $\{C\}_\phi$ constitute the sector labelled by the null irreducible string: $I \equiv \phi$. For this sector, $D_\phi \sim (27/4)^{L/3}$.

(vi) Consider picking a configuration C_{ran} totally at random from amongst the 2^L possible configurations — an important case, since most configurations are of this type. We may ask for the likely length l_{ran} of the corresponding IS. We find that $l_{\text{ran}} \approx L(1 - 3\mu^{-2}/2)$.

5. Trimers: Dynamics

We have seen that the DE process partitions phase space into an exponentially large number of sectors. A strong variation in the behaviour of $C_i(t)$, from one sector to another, is found in Monte Carlo studies [4]. For instance, in the sector with IS $= (\uparrow\uparrow\downarrow)^{fL}$ [i.e. the string $\uparrow\uparrow\downarrow\uparrow\uparrow\downarrow\cdots$ formed by repeating $\uparrow\uparrow\downarrow fL$ times] with $f < 1/3$, we find $C_i(t) \sim 1/t^{1/2}$. In the sector with IS $= (\uparrow\downarrow)^{fL}$ with $f < 1/2$, we find $C_i(t) \sim \exp(-bt^{1/2})$. If we start with a totally random initial configuration, we find that $C_i(t) \sim 1/t^{1/4}$. In the sector where the IS is a periodic string $(\uparrow\uparrow\downarrow\downarrow\uparrow\downarrow)^{fL}$ with $f < 1/6$, the autocorrelation function $C_i(t)$ depends on which sublattice site i lies (in a 3-sublattice decomposition of the lattice). $C_i(t) \sim 1/t^{1/2}$ for i on two sublattices, while $C_i(t) \sim \exp(-bt^{1/2})$ for i on the third sublattice.

An explanation of dynamical diversity can be found in terms of the IS: besides being a label on sectors, it has a dynamical meaning as discussed below. The point is that elements of the IS are themselves a subset of the full configuration of spins — namely, those that are left undeleted at the end of repeated application of the triplet reduction procedure (the 'deletion do loop' referred to above). Under time evolution, the configuration evolves, and as a result, the IS elements move on the lattice, though keeping their relative ordering intact (as the IS is a constant of the motion). In other words, the elements of the IS behave like a set of hard-core random walkers with spin (HCRWS), random-walking through the background of reducible spins. The diffusive motion of the HCRWS constitute the slow modes of the original model. We thus expect the HCRWS dynamics to account for the long-time properties of the deposition-evaporation system [4].

If we ignore the reducible background fluctuations (which occur rapidly, and contribute very little to the long-time behaviour), quite a lot can be said about HCRWS dynamics, as a good deal is known already about the diffusion of particles on a lattice,

with mutual hard-core interactions (the exclusion process) [10]. The arrangement of spins on the particles in no way affects the particle dynamics, but is of course crucial in determining the spin autocorrelation function, which is what we need. The latter involves a convolution over the spin pattern (which in turn identifies a particular sector of the deposition-evaporation model), and the conditional probability in the exclusion process that particle m is at a particular site at $t = 0$, and particle $m + \Delta m$ is at the same site at time t.

The HCRWS model explains the variety of long-time behaviours found in different sectors in which the density of walkers l/L is finite, as detailed in the first paragraph of this section. However, in the null-IS sector, consisting of the set of all fully reducible configurations $\{C\}_\phi$, the dynamics is different, as it is the reducibility fluctuations which matter (these were the fast modes in the other sectors, and could be ignored there as a result). Numerical data from Monte Carlo simulations [4], and exact diagonalization of the transfer matrix for finite size systems [11] suggests that in this case fluctuations are slowly decaying and sub-diffusive in nature. The estimated value of the dynamical exponent $z \approx 2.5$ is quite different from the value for the Kardar–Parisi–Zhang type of stochastic growth models [12] which in one dimension are known to have $z = 3/2$.

6. Conclusion

Models of k-mer deposition and evaporation provide examples of many-state systems about which we can say quite a lot. Central to the understanding of both the phase space decomposition into exponentially many sectors, and also the diversity of asymptotic dynamical behaviour, is the irreducible string. The IS provides a compact and convenient means of labelling disjoint sectors, and thus provides a solution to the 'assignment problem' posed in the Introduction. Besides, the elements of the IS have a dynamical meaning: their diffusive movement through the lattice (with an implied hard core constraint between different elements) constitute the slow modes of the system; this allows for a characterization of the asymptotic decay laws of the autocorrelation function, and an understanding of their sector-wise variation.

Although we have not discussed it in this paper, the IS in fact encapsulates an infinite number of conservation laws. The way to construct conserved quantities (all of which are mutually commuting functions of $\{\sigma_1^z, \sigma_2^z \cdots \sigma_L^z\}$) is discussed in [3]. It is entirely possible that there is yet another family of conservation laws, which involve transverse components of the spin, σ_i^+ and σ_i^-. If they do exist, these transverse conservation laws would provide direct connections between the properties of different sectors. For dimers, where the decomposition of phase space is not nearly as severe, we know that transverse conservation laws do exist — they correspond to the conservation of total spin in that case. The elucidation of analogous conservation laws in many-state systems remains a challenging open problem.

Acknowledgements

It is a pleasure to acknowledge fruitful collaborations on deposition-evaporation systems with Robin Stinchcombe, Marcelo Grynberg and Deepak Dhar.

References

[1] M Barma, M D Grynberg and R B Stinchcombe, *Phys. Rev. Lett.* **70**, 1033 (1993)
[2] R B Stinchcombe, M D Grynberg and M Barma, *Phys. Rev.* **E47**, 4018 (1993)
[3] D Dhar and M Barma, *Pramana – J. Phys.* **41**, L193 (1993)
[4] M Barma and D Dhar, *Phys. Rev. Lett.* **73**, 2135 (1994)
[5] P J Flory, *J. Am. Chem. Soc.* **61**, 1518 (1939)
[6] J J Gonzalez, P C Hemmer and J S Høye, *Chem. Phys.* **3**, 228 (1974)
[7] J W Evans, *Rev. Mod. Phys.* **65**, 1281 (1993)
[8] T Moriya in *Magnetism* edited by G T Rado and H Suhl (Academic Press, New York, 1963) vol. 1
[9] M Barma and D Dhar, in *Proceedings of the International Colloquium on Modern Quantum Field Theory II* edited by S R Das, A Dhar, G Mandal and S R Wadia (World Scientific, Singapore, to appear)
[10] T M Liggett, *Interacting Particle Systems* (Springer, New York, 1985)
[11] P B Thomas, M K Hari and D Dhar, *J. Phys.* **A27**, L831 (1994)
[12] M Kardar, G Parisi and Y Zhang, *Phys. Rev. Lett.* **56**, 889 (1986)
J Krug and H Spohn in *Solids far from equilibrium* edited by C Godrèche (Cambridge Univ. Press, Cambridge, 1991)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 2,
August 1997
pp. 161–192

Cosmological *N*-body simulations

J S BAGLA and T PADMANABHAN
Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India

MS received 14 November 1996; revised 24 April 1997

Abstract. In this review we discuss cosmological *N*-body codes with a special emphasis on particle mesh codes. We present the mathematical model for each component of *N*-body codes. We compare alternative methods for computing each quantity by calculating errors for each of the components. We suggest an optimum set of components that can be combined to reduce the overall errors in *N*-body codes.

Keywords. Galaxy formation; cosmology; numerical simulations.

PACS Nos 02.00; 04.00; 04.25; 98.62; 98.80

1. Introduction

Observations suggest that the present universe is populated by very large structures like galaxies, clusters of galaxies etc. Current models for formation of these structures are based on the assumption that gravitational amplification of small perturbations leads to the formation of these large scale structures. In the absence of analytical methods for computing quantities of interest, numerical simulations are the only tools available for the study of clustering in the non-linear regime. The last two decades have seen a rapid development of techniques and computing power for cosmological simulations and the results of these simulations have provided valuable insight into the study of structure formation. In this paper we will describe some aspects of the cosmological *N*-body simulation, based on the code developed and used by the authors. We will stress the key physical and numerical ideas and detail some useful tests of the *N*-body code. After an initial introduction to cosmological *N*-body codes, the discussion of *N*-body codes will focus on particle mesh (PM) codes. (For a comprehensive discussion of numerical simulations using particles, see [1]).

An *N*-body code consists of two basic modules: one part computes the force field for a given configuration of particles and the other moves the particles in this force field. These two are called at each step to ensure that the force field and the particle trajectories evolve in a self consistent manner. Apart from these we also need to set up the initial conditions and write the output. Thus the basic plan of an *N*-body code follows the flow chart shown in figure 1.

Structure of these modules depends on the specific application at hand. Here we outline some features that are particular to cosmological *N*-body codes. To make quantitative

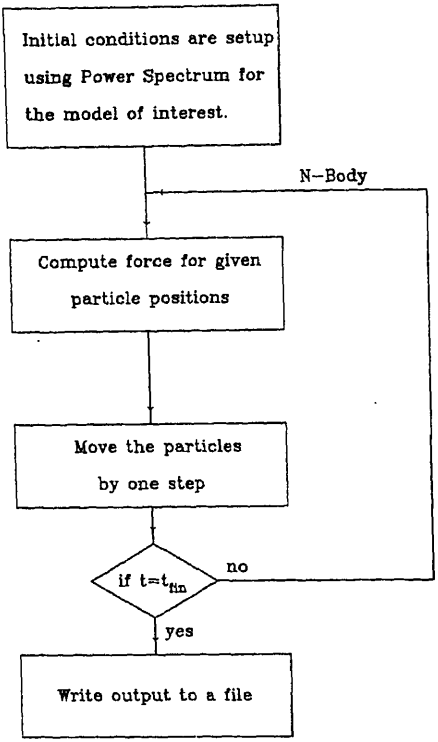

Initial conditions are setup using Power Spectrum for the model of interest.
N-Body
Compute force for given particle positions
Move the particles by one step
if t=t_fin
no
yes
Write output to a file

Figure 1. Flow chart for an N-body code.

statements about the required parameters of an N-body simulation we shall assume that the particles populate a region of volume V. We shall also assume that we are using N_p particles to describe the density field. The following physical requirements have to be taken into account while writing cosmological N-body codes.

- The universe is filled with matter over scales that are much larger than the scales of interest in an N-body simulation. Density averaged over larger and larger scales in the universe approaches a constant value. Thus the simulation volume V cannot be assumed to exist in isolation and we must fill the region outside this volume by some method. Periodic boundary conditions are the only viable solution to this problem. In absence of periodic boundary conditions most of the matter in the simulation volume tend to collapse towards the centre of the box, giving rise to a spurious, large rate of growth for perturbations. (A compromise solution, called quasiperiodic boundary conditions, has been tried to solve this problem for tree codes. In this scheme periodic boundary conditions are used to restructure the simulation volume to bring a particle at the centre of the box while computing force at its position. This is done for each particle so that there is no strong tendency to collapse towards the centre of the box [2]). The most natural geometry for a volume with periodic boundary conditions is a cube.
- We would like the evolution of perturbations to be independent of the specific boundary conditions. Thus the simulation volume should not be dominated by any one

its own periodic copies will influence its later evolution).

We require the average density in the box to equal the average density of the universe. Thus, perturbations averaged at the scale of the box must be ignorable at all times for the model of interest, i.e., $\sigma(R = V^{1/3}) \ll 1$. [Here $\sigma(R)$ is the rms dispersion in the mass averaged at scale R.] For example, in case of the standard CDM model this would require the box to be at least $100\,h^{-1}$ Mpc [at $z = 0$] in extent.

- We would like to probe scales that are sufficiently non-linear in order to justify the use of N-body simulations. If we wish to compare the results with the real universe then we should be able to study the formation of structures with a large range in scales. In other words, the mass of individual particles in an ideal simulation must be less than the mass of the smallest structure of interest. Therefore, the smallest number of particles required to cover the relevant range of scales to study galaxy clustering is

$$N \geq \frac{M(100\,h^{-1}\,\mathrm{Mpc})}{M_{\mathrm{galaxy}}} \simeq \frac{5 \times 10^{18} M_{\odot}}{10^{11} M_{\odot}} \simeq 5 \times 10^{7}. \tag{1}$$

  We need a very large number of particles to represent the density field over this range of scales. Therefore, most numerical techniques used in cosmological N-body codes are oriented towards reducing the number of operations and stored quantities per particle.
- We approximate the collection of a very large number of particles in the universe by one particle in an N-body simulation. Therefore the particles in an N-body simulation must interact in a purely collisionless manner.

The periodic boundary conditions and a very large number of particles are the key considerations in developing the detailed algorithm for an N-body code. Of the two main components in an N-body code, integration of the equation of motion is an algorithm of order $O(N_p)$. The calculation of force, if done by summing the force over all pairs, is a process of order $O(N_p^2)$. Therefore, the calculation of force is likely to be more time consuming than evolving the trajectories in N-body codes with a large number of particles. It can be shown that, for particle numbers greater than a few hundred, direct computation of force takes an excessively long time even on the fastest computers. (The very high speed special purpose computers are an exception to this [3].) Three schemes have been evolved to address this problem and replace direct summation over pairs by methods that are less time consuming. These are

- *Particle mesh (PM)*: Poisson equation is solved in the Fourier domain and the potential force is computed on a fixed grid. The force is then interpolated to particle positions for moving the particles. Density field, the source for gravitational potential, is also computed on the same mesh/grid from particle positions by using a suitable interpolating function. The 'smoothing' of particles limits the resolution of such simulations but ensures collisionless evolution. (See [4] and [5] for an excellent discussion of PM codes).

  Particle–particle particle mesh (P^3M): This scheme [6] improves the resolution of PM method by adding a correction to the mesh force for pairs of particles with separation of the order of, and smaller than the grid length. The number of operations required for

this correction is proportional to $N_p n(R)$ where $n(R)$ is the average number of neighbouring particles within a distance R. This can be written as $N_p \bar{n}(1+\bar{\xi}(R))$, where $\bar{n}$ is the average number density and $\bar{\xi}$ is the averaged correlation function. It is obvious that such corrections can become time consuming in highly clustered situations [when $\bar{\xi}(R) \gg 1$].

- *Tree*: In this scheme, information about the density field is set up in the form of a hierarchical tree. Each level of the tree specifies position of the centre of mass and the total mass for a region in the simulation volume. Force from particles in distant regions in the simulation box is approximated by the force from the centre of mass for particles in that region. This leads to a reduction in the number of operations required for calculating force [7, 2]. To estimate the number of operations required for setting up the tree structure and evaluate the force, let us assume that the simulation volume is a cube and is subdivided into eight equal parts at each level. This subdivision of cells is continued till we have at most one particle per cell at the smallest level. Each particle is parsed at most once at each level, therefore the upper bound on the total number of operations is proportional to $N_p \ln (L_{\rm box}/l_{\rm min})$ where $l_{\rm min}$ is the smallest inter-particle separation for the given distribution of particles. We have

$$l_{\rm min} \simeq n_{\rm max}^{-1/3} = \delta_{\rm max}^{-1/3} \bar{n}^{-1/3} = N_p^{-1/3} L_{\rm box} \delta_{\rm max}^{-1/3}, \tag{2}$$

where $\bar{n}$ is the average number density, $\delta_{\rm max}$ is the maximum density contrast and $n_{\rm max}$ is the highest number density in the given distribution of particles. This implies that the upper bound on number of operations is proportional to $O(N_p \ln N_p \delta_{\rm max})$. Incorporating periodic boundary conditions is a very nontrivial problem in the context of tree codes. (See [8] for a discussion of periodic boundary conditions for tree codes.)

Amongst these methods the PM scheme has two advantages over the other methods. Firstly it ensures a collisionless evolution (see [9]; [10]). Secondly it has the simplest algorithm and it is also the fastest of the three methods discussed above. In the remaining discussion we shall focus only on PM codes. However, apart from computation of force, all other components are the same in all these codes (with some minor variations) and most of the conclusions about relative merits of the available alternatives are applicable for all three types of codes.

2. Moving the particles

In this section we will first present the system of equations that describe the evolution of trajectories of particles. This will be followed by a description of the methods used for numerical integration of equation of motion.

2.1 *Equation of motion*

It can be shown that the evolution of perturbations in a non-relativistic medium in an expanding background can be studied in the Newtonian limit at scales that are much smaller than the Hubble radius $d_H = cH_0^{-1}$. The equations for a set of particles interacting

only through the gravitational force can be written as

$$\ddot{\mathbf{x}} + 2\frac{\dot{a}}{a}\dot{\mathbf{x}} = -\frac{1}{a^2}\nabla_x \varphi$$

$$\nabla_x^2 \varphi = 4\pi G a^2 \bar{\varrho}\delta = \frac{3}{2} H_0^2 \Omega_0 \frac{\delta}{a}. \tag{3}$$

Here the last equality follows from the Friedmann equations that describe evolution of the scale factor for the universe. The variables used here are the comoving co-ordinates x, time t, scale factor a, gravitational potential due to perturbations φ, density ϱ and density contrast δ. Cosmological parameters used in this equation are the Hubble's constant H_0 and the density parameter contributed by non-relativistic matter Ω_0.

N-Body simulations integrate this equation numerically for each particle. Numerical integration can be simplified by modifying the equation of motion so that the velocity dependent term is removed from the equation of motion. This can be achieved by using a different time parameter θ [11]. This is defined as

$$\mathrm{d}\theta = H_0 \frac{\mathrm{d}t}{a^2}, \tag{4}$$

in terms of which we have

$$\frac{\mathrm{d}^2 \mathbf{x}}{\mathrm{d}\theta^2} = -\frac{3}{2}\Omega_0 a^2 \nabla \chi,$$

$$\nabla^2 \chi = \frac{\delta}{a}. \tag{5}$$

In this form, all the cosmological factors can be clubbed in the source term in the equation of motion, making numerical implementation much simpler.

2.2 *Numerical integration*

In any N-body code, a great deal of computer time is devoted to integration of the equation of motion. The basic idea of time integration is simple: the equation of motion expresses the second derivative of position in terms of position, velocity and time. Time integration translates this into the subsequent changes in position and velocity. In the following discussion we will develop the idea of numerical integration of trajectories.

Writing the position and velocity at time $t + \epsilon$ in a Taylor series about the position and velocity at time t, we have

$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t) + \tfrac{1}{2}\epsilon^2 a_i(t) + \cdots,$$

$$v_i(t+\epsilon) = v_i(t) + \epsilon a_i(t) + \tfrac{1}{2}\epsilon^2 j_i(t) + \cdots. \tag{6}$$

Here x is the position, v is the velocity, a is the acceleration and j is the rate of change of acceleration, and the subscript is used to identify the particle. We can use the equation of motion to express acceleration and the higher derivatives of position in terms of positions and velocities. Therefore, in principle, we can find the new position and velocity with

arbitrary accuracy. However, such a scheme is not very practical from the computational point of view as it requires an infinite series to be summed. A more practical method consists of truncating the series at some point and using sufficiently small time steps to ensure the required level of accuracy. To illustrate this point, let us consider terms up to first order in the above series, we get

$$v_i(t+\epsilon) = v_i(t) + \epsilon a_i(t) + O(\epsilon^2),$$
$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t) + O(\epsilon^2). \tag{7}$$

This is the Euler's method of solving differential equations and the error in the solution is of order ϵ^2. Thus choosing a smaller time step leads to smaller error in each step. However, the number of steps required to integrate over a given interval in time is inversely proportional to the time step so that the overall error changes only linearly with step size.

Let us consider terms up to ϵ^2 in order to device a more accurate integrator.

$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t) + \tfrac{1}{2}\epsilon^2 a_i(t) + O(\epsilon^3),$$
$$v_i(t+\epsilon) = v_i(t) + \epsilon a_i(t) + \tfrac{1}{2}\epsilon^2 j_i(t) + O(\epsilon^3). \tag{8}$$

This equation contains the rate of change of acceleration and therefore is not very useful in this form. To simplify these equations without losing accuracy, let us specify j as

$$j(t) = \frac{a(t+\epsilon) - a(t)}{\epsilon} + O(\epsilon). \tag{9}$$

This reduces the above equations to the following form

$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t) + \tfrac{1}{2}\epsilon^2 a_i(t) + O(\epsilon^3),$$
$$v_i(t+\epsilon) = v_i(t) + \frac{\epsilon}{2}(a_i(t) + a_i(t+\epsilon)) + O(\epsilon^3). \tag{10}$$

This method can be used for velocity independent forces and is identical to the Leap–Frog method. The standard Leap–Frog method involves updating the velocity and position at an offset of half a step. To see the equivalence of the above equations and the Leap–Frog method let us consider the expressions for velocity at $t + \epsilon/2$ and $t - \epsilon/2$.

$$v_i(t+\epsilon/2) = v_i(t) + \frac{\epsilon}{2}a_i(t) + \frac{1}{8}\epsilon^2 j_i(t) + O(\epsilon^3),$$
$$v_i(t-\epsilon/2) = v_i(t) - \frac{\epsilon}{2}a_i(t) + \frac{1}{8}\epsilon^2 j_i(t) + O(\epsilon^3). \tag{11}$$

These two equations can be combined to give

$$v_i(t+\epsilon/2) = v_i(t-\epsilon/2) + \epsilon a_i(t) + O(\epsilon^3). \tag{12}$$

We can also use the expression for velocity at $(t+\epsilon)$ to write

$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t+\epsilon/2) + O(\epsilon^3). \tag{13}$$

These two equations can now be combined to give the Leap–Frog method

$$x_i(t+\epsilon) = x_i(t) + \epsilon v_i(t+\epsilon/2) + O(\epsilon^3),$$
$$v_i(t+3\epsilon/2) = v_i(t+\epsilon/2) + \epsilon a_i(t) + O(\epsilon^3). \tag{14}$$

This is called the Leap–Frog method as it updates velocities halfway between the step that is used to update the position. For velocity dependent forces these methods have to be modified into a predictor–corrector form.

It is possible to improve the accuracy by using integrators that retain more terms in the Taylor series expansion. However most of these schemes require many evaluations of force for each step making them computationally uneconomical.

2.3 *Choosing the time step*

The time step for integrating the equation of motion can be chosen in many ways. In all of these methods one defines a suitable criterion (like energy conservation) and then obtains the largest time step that will satisfy the requirement. Following is a list of possible criteria:

- *Monitoring validity of Irvine–Layzer equation.* This is done by integrating the Irvine–Layzer equation [12, 13, 14]

$$T + U + \int \frac{da}{a}(2T + U) = C \tag{15}$$

  and monitoring the constant C. Variation in C with respect to total energy $T + U$ can be used as an indicator of non conservation. (This is normally done with a different form of this equation so that only the kinetic or the potential energy term contributes to the integral [6]). In schemes where the force is computed at mesh points and then interpolated to particle positions the force at the particle position, $\mathbf{f}(\mathbf{x}) \neq -\nabla_{\mathbf{x}}\varphi$, i.e., the force does not equal the gradient of potential at a point. This leads to an apparent non conservation of energy. If we include a correction term in the energy equation to account for this fact [6] then it can be shown that it is possible to improve conservation of energy. This correction involves a direct sum over all pairs of particles and is therefore an almost impossible task to implement for a simulation with a large number of particles. Thus this is not a practical method of deciding the time step.
- *Convergence of final positions and velocities.* If we evolve the trajectories of a set of particles with different time steps, then we expect the final positions and velocities to converge towards the correct value as we reduce the time step. This can be used to decide the time step.
- *Reproducibility of initial conditions.* This is a more stringent version of the test outlined above. This method ensures that the solution we get is correct and errors are not building up systematically. If we run the particles forward and then back again we should in principle get back the initial positions. Although we ignore the decaying mode while integrating trajectories forward in time, it tends to dominate if we evolve the system in the opposite direction. We can overcome this difficulty by not evolving the potential during this test.

Table 1. This table lists the values of ϵ used in forward and backward integration of trajectories of a set of particles in an external potential and the corresponding error in recovering the initial conditions. Here we have used one grid length as the unit of length.

ϵ	$(\Delta x)_{\rm rms}$
0.1	5.4×10^{-4}
0.05	7.3×10^{-5}
0.02	3.5×10^{-5}
0.01	1.2×10^{-5}
0.005	5.3×10^{-6}

We find the last test to be the most stringent one and we use it to fix the time step for an arbitrary potential by the following construct. Consider an arbitrary potential ψ which has the maximum value of the magnitude of its gradient as $g_{\rm max} = |\nabla\psi|_{\rm max}$. We can write the equation of motion eq. (5) for a small interval $\Delta\theta$ as

$$\frac{\epsilon}{(\Delta\theta)^2} = \frac{24}{\theta^4}\, g_{\rm max}. \tag{16}$$

Here we have used ϵ to represent the largest displacement in time $\Delta\theta$. We have specialized to $\Omega_0 = 1$ in writing the above equation. In which case, $\theta = -2/\sqrt{a}$. For a given value of ϵ we can fix the time step as

$$\Delta\theta = \theta^2 \sqrt{\frac{\epsilon}{24 g_{\rm max}}}. \tag{17}$$

We can use the test mentioned above to fix the value of ϵ. Table 1 lists a few values of ϵ and $(\Delta x)_{\rm rms}$ for a CDM simulation [box size $= 64\,{\rm h}^{-1}$ Mpc] with 64^3 particles in a 64^3 box. The system was evolved from $a = 0.02$ to 0.25 and then back to $a = 0.02$. We use $\epsilon = 0.005$ for our simulations as we find the error in this case to be acceptable.

3. Calculating the force

An N-body code solves the equation of motion and the Poisson equation in a self-consistent manner. Therefore, moving the particles and computing the force for a given distribution of particles are the two most important components of an N-body code. In this section we will discuss computation of force at mesh points for a given distribution of particles. For simplicity we shall assume that all the particles have the same mass. Generalization to particles with different masses is not difficult but can strain the resources like RAM. (Setting up initial conditions for simulations that have particles with different masses is much simpler than setting up initial conditions for simulations that use particles with equal mass. See § 4 for details.)

In particle mesh codes the force at the particle position can be computed from the potential defined at mesh points in two ways. These are:

Energy conserving schemes: The force is computed at particle positions by using the gradient of the interpolating function

$$\mathbf{f}(\mathbf{x}) = -\nabla_{\mathbf{x}} \left(\sum_{\mathbf{x}_i} W(\mathbf{x}, \mathbf{x}_i) \psi(\mathbf{x}_i) \right). \tag{18}$$

The sum is over all the mesh points. An additional requirement is that the same interpolating function W should be used to compute the density field on the mesh from the distribution of particles. These are called energy conserving schemes as the force and gravitational potential are related to each other as $\mathbf{f}(\mathbf{x}) = -\nabla_{\mathbf{x}} \psi(\mathbf{x})$.

- *Momentum conserving schemes*: The force is computed on the mesh and then interpolated to the particle position.

$$\mathbf{f}(\mathbf{x}) = -\sum_{\mathbf{x}_i} W(\mathbf{x}, \mathbf{x}_i) \nabla_{\mathbf{x}_i} \psi(\mathbf{x}_i). \tag{19}$$

Momentum conserving schemes also require the same interpolating function to be used for computing density field on the mesh. It can be shown that in these schemes the force due to the ith particle on the jth particle is same as the force due to jth particle on the ith particle ($\mathbf{f}_{ij} = -\mathbf{f}_{ji}$). This clearly is a necessary condition for momentum conservation.

Use of mesh leads to anisotropy in force at scales comparable to the mesh size. A simple method for limiting these anisotropies is to choose a configuration that leads to a softer force at the mesh scale. Collisionless evolution also requires the force to drop rapidly below the average inter-particle separation [9]. We choose to work with the momentum conserving schemes as these lead to a softer force at small scales in comparison with the energy conserving schemes.

The algorithm for computing force at particle positions involves the following major steps.

Computing the density field: Mass of particles is assigned to mesh points by using an interpolating function. Density contrast is calculated from this 'mass field' by using the definition

$$\delta = \frac{M}{\bar{M}} - 1. \tag{20}$$

Here $\bar{M}$ is the average mass per mesh point. Fast Fourier transform technique is used to compute the Fourier components of density contrast.

Solving the Poisson equation: The Poisson equation is solved in Fourier space. The FFT can be used to compute the gravitational potential on the mesh.

- *Computing the force*: The gravitational force equals the negative gradient of the gravitational potential, therefore the next step in algorithm is computing the gradient of potential. This can be done either in Fourier space or in real space by using some scheme for numerical differentiation. Force at the particle positions is computed by interpolation from the mesh points.

In the following subsections we will discuss each of these steps in greater detail.

3.1 *Computing the density contrast*

Smoothing [interpolating] functions are used to assign mass to the lattice points for a given configuration of particles. This is done by computing the following sum for every mesh point.

$$M(\mathbf{x}_i) = \sum_{\text{all particles}} M_{\text{particle}} W(\mathbf{x}, \mathbf{x}_i). \tag{21}$$

Here $\mathbf{x}_i$ are the co-ordinates of mesh points, $\mathbf{x}$ are the coordinates of particles and W is the interpolating function. As we have assigned same mass to all the particles we can scale the mass at mesh sites with the particle mass.

Density contrast is computed from the mass defined at mesh points by using eq. (20). FFT is then used to compute its Fourier transform.

The smoothing function W should satisfy certain algebraic, computational and physical requirements.

- The sum of weights given to lattice points should equal unity. The interpolating function should be continuous and preferably differentiable as well.
- The smoothing function should be peaked at small separations and should decay rapidly at large separations. The smoothing function should have the smallest possible base, i.e., the number of lattice separations over which it is non-zero should be small. The number of arithmetic operations required for interpolation is proportional to the cube of the number of lattice cells over which the interpolating function is non-zero.
- The interpolating function should be isotropic.

It is clear that we cannot satisfy all these criteria fully and at the same time and therefore we have to consider a compromise solution. The first simplification that is made is to break-up the three dimensional interpolating function into a product of three one dimensional interpolating functions. This reduces the complexity of the problem but also implies that the interpolating function cannot be manifestly isotropic.

These considerations, and the above mentioned simplifying assumption then leads to the following set of interpolating functions.

- *Nearest grid point (NGP)*: This is the simplest interpolating function. In this assignment scheme all the mass is assigned to the nearest grid point

$$W(x, x_i) = \begin{cases} 1 & (\text{for } |x - x_i| \leq L/2) \\ 0 & (\text{for } |x - x_i| > L/2), \end{cases} \tag{22}$$

  where L is the grid spacing. This clearly satisfies the algebraic condition that the sum of weights should equal unity. However, this function is neither continuous nor differentiable. The leading term in the error in force of a point mass computed using this mass assignment has a dipole-like behaviour.
- *Cloud in cell (CIC)*: This is the most commonly used interpolating function. Linear weights are used and mass is assigned to the two nearest grid points. The function is continuous but not differentiable.

$$W(x,x_i) = \begin{cases} 1 - |x - x_i|/L & (\text{for } |x - x_i| \leq L) \\ 0 & (\text{for } |x - x_i| > L). \end{cases} \tag{23}$$

Triangular shaped cloud (TSC): The third function we wish to discuss is a quadratic spline that is continuous and also has a continuous first derivative. In this case the mass is distributed over three lattice points in each direction

$$W(x,x_i) = \begin{cases} \frac{3}{4} - \left(\frac{\Delta x}{L}\right)^2 & (\text{for } \Delta x \leq L/2) \\ \frac{9}{8}\left(1 - \left|\frac{2\Delta x}{3L}\right|\right)^2 & (\text{for } L/2 \leq \Delta x \leq 3L/2) \\ 0 & (\text{for } 3L/2 \leq \Delta x) \end{cases} \tag{24}$$

Here $\Delta x = |x - x_i|$ is the magnitude of separation between the particle and the mesh point. This function satisfies the requirements from the physical point of view but it is generally considered expensive from the computational point of view.

It is obvious that these smoothing functions become more and more computationally intensive as we go for a larger base. All the functions described above satisfy the essential algebraic requirements, so that the final choice is a compromise between the accuracy of the force field obtained using these and the computational requirements. However, we do not go beyond TSC as the base becomes very large and the transverse error in force becomes unacceptably large.

Figure 2 compares these smoothing functions in a realistic situation. We have plotted the average weight assigned to a point that is at a distance r from the mesh point of interest. The curve in each of the panels shows the average weight where the average is computed over all directions $[\theta, \phi]$ while keeping r fixed. The vertical bars show the rms dispersion about this average to depict the level of anisotropies in each of these functions. It is clear that TSC is the most isotropic of these functions and CIC is a 'reasonable' compromise between isotropy and computational requirements. We will use the TSC smoothing function in the N-body code for all applications.

3.2 *Solving the Poisson equation*

The equation we wish to solve is

$$\nabla^2 \psi = \frac{\delta}{a}. \tag{25}$$

This equation is to be solved with periodic boundary conditions. In particle mesh codes the equation is solved in the Fourier space. To solve this equation with periodic boundary conditions we must ensure that only the harmonics of the fundamental wave number $k_f = 2\pi/NL$ contribute to the function. (NL is the size of the simulation box.) This type of sampling of k space is also required by the fast Fourier transforms. Therefore as long as we use FFT for computing Fourier transforms the correct boundary conditions are automatically incorporated.

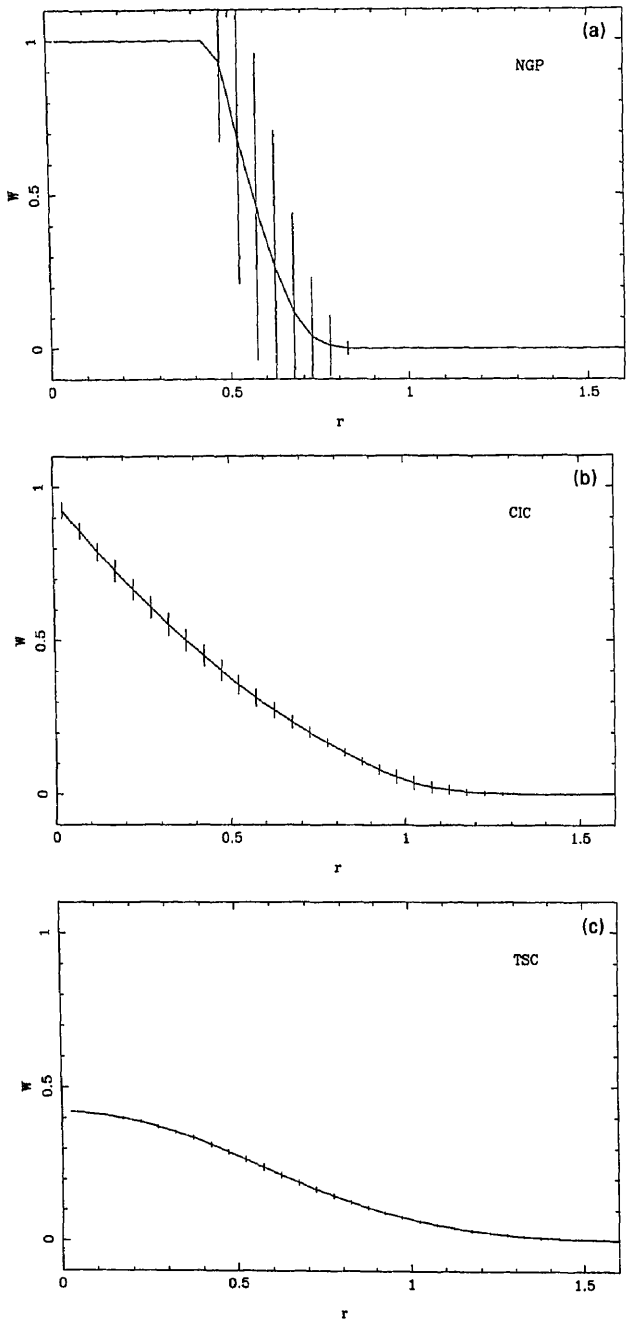


Figure 2. This figure shows the weight assigned to a point that is at a distance r from the mesh point of interest. The curve in each of the panels shows the average weight where the average is computed over all directions $[\theta, \phi]$, keeping r fixed. The vertical bars show the rms dispersion about this average to depict the level of anisotropies in these functions.

Two types of Green's functions are commonly used for solving the Poisson equation in Fourier space. These are

- *Poor man's Poisson solver*: This method uses the continuum Green's function

$$\psi_k = \frac{\delta_k}{a} G(k),$$
$$G(k) = -\frac{1}{k^2}. \tag{26}$$

The main advantage of this is that it is the true Green's function. Also, it is isotropic by construction.

Periodic Green's function: An alternative is to substitute Fourier transform of a discrete realization of the Laplacian. A Green's function derived in this manner tends to increase power at small scales and leads to larger anisotropies in the force field. A commonly used Green's function of this class is

$$G(k) = -\frac{L^2}{4[\sin^2 k_x L/2 + \sin^2 k_y L/2 + \sin^2 k_z L/2]}. \tag{27}$$

Here L is the grid spacing and equals unity in the units that we are using here. This Green's function is invariant under the transformation $k_x \rightarrow k_x + 2\pi$. All FFT routines use such a transformation to rearrange wave modes for ease of manipulation (see § 3). Therefore this Green's function is easier to implement as compared to the continuum version as one does not have to worry about the rearrangement of wave modes. The main disadvantage of this Green's function is that it is anisotropic at small scales.

We have plotted the two Green's functions in figure 3. Like figure 2 the curve here shows the average over angular coordinates for any value of $k = |\mathrm{k}|$. The vertical bars show the rms dispersion about this average and are an indicator of anisotropy. It is clear from this figure that the periodic Green's function enhances power at small scales, beyond the true value, and also introduces some anisotropy in the potential.

We compared these two Green's functions by solving the Poisson equation for the density contrast

$$\delta = A\left(\frac{r^2}{\sigma^4} - \frac{3}{\sigma^2}\right) \exp\left[-\frac{r^2}{2\sigma^2}\right]. \tag{28}$$

The solution to the Poisson equation for this source can be obtained analytically and the potential is

$$\psi = \frac{A}{a} \exp\left[-\frac{r^2}{2\sigma^2}\right] + \text{const.} \tag{29}$$

We solved the Poisson equation numerically using the two Green's functions defined above on a 64^3 mesh. We computed the rms deviation from the true solution in each case for a large range of values for σ. The result of this analysis is presented in figure 4 that shows the error as a function of σ. This curve clearly shows that the periodic kernel introduces more error than the poor man's Poisson solver.

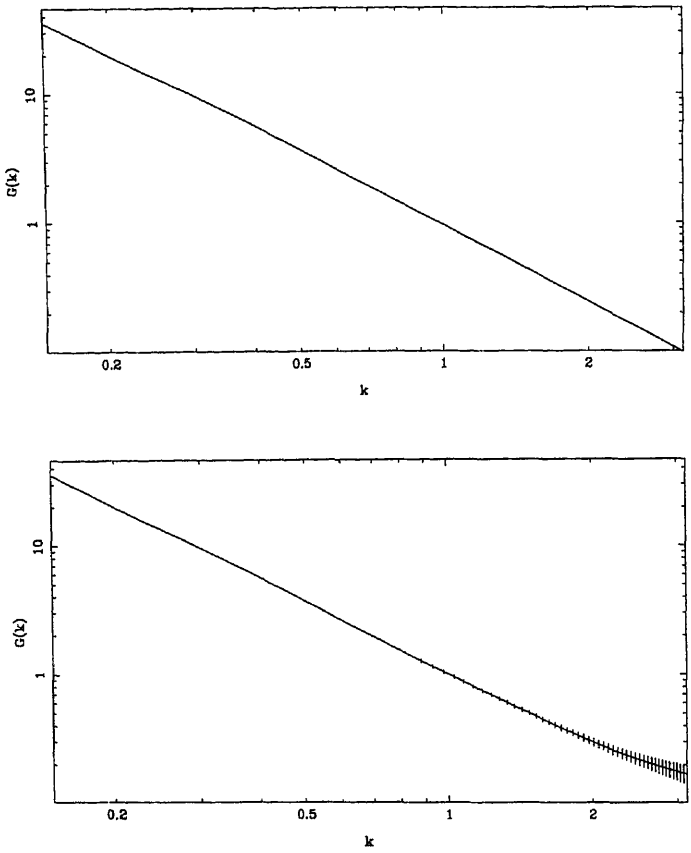


Figure 3. This figure shows the two Green's functions described in the text. Upper panel shows the continuum Green's function and the lower panel shows the Green's function derived from discrete representation of the Laplacian. Curves in each panel show the average of $G(k)$ over angular coordinates for any value of $k = |\mathbf{k}|$. The vertical bars show the rms dispersion about this average and are an indicator of anisotropy.

3.3 *Computing the gradient of gravitational potential*

Numerical differentiation is used to compute the force $[-\nabla\psi]$ on the mesh. The gravitational potential ψ is also defined on the mesh. Interpolation is used to compute the force at the particle position.

Here we will discuss three most commonly used methods for computing the gradient of potential in N-body codes.

- The most accurate and fastest method for computing derivatives uses fast Fourier transforms. The Fourier components of force are computed directly from the Fourier components of gravitational potential. Only possible source of errors here is the FFT routine itself. The FFT routine used here gives an error of about 10^{-4} % for Fourier

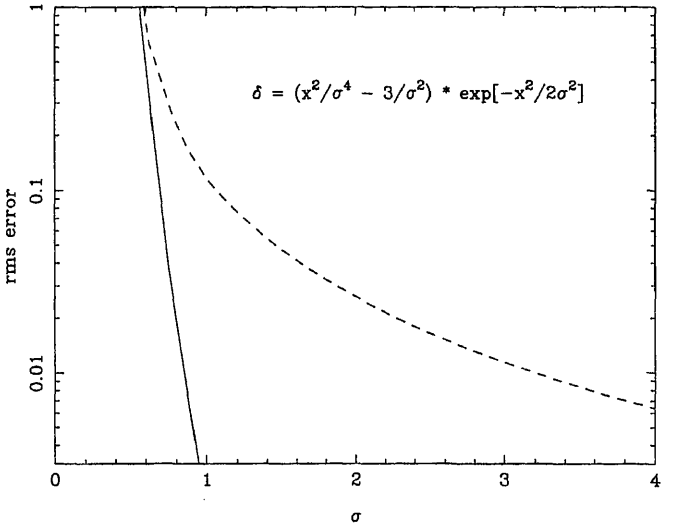


Figure 4. This figure shows the rms error in solving Poisson equation as a function of the parameter σ defined in the text. The dashed line shows the error for the periodic kernel whereas the thick line shows the error for the poor man's Poisson solver. The comparison was carried out on the grid so as to avoid mixing errors in solving the equation with errors introduced by interpolation etc.

transforming a sine wave in a 64^3 box. (This error was estimated by comparing numerical values of Fourier components with their expected value.)

The expression for Fourier transform of a derivative is

$$f'_k = -ikf_k. \tag{30}$$

The simplest method for computing the derivative in real space is the mid point rule.

$$\frac{\partial f}{\partial x} = \frac{f(x+h) - f(x-h)}{2h} + O(h^2). \tag{31}$$

The equivalent expression for computing the derivative on the mesh is

$$\frac{\partial f}{\partial n} = \frac{f(n+1) - f(n-1)}{2}. \tag{32}$$

For an individual sine wave, the fractional error varies as k^2 for small k. Here, as for every operation in cosmological simulations, periodic boundary conditions are used to compute the derivative near the edges.

It is possible to construct a more accurate formula by using the method of undetermined coefficients (see for example [15]). A five point formula for differentiation that gives fractional errors proportional to k^4 for small k is

$$\frac{\partial f}{\partial x} = \frac{f(x-2h) - f(x+2h) + 8(f(x+h) - f(x-h))}{12h} + O(h^5). \tag{33}$$

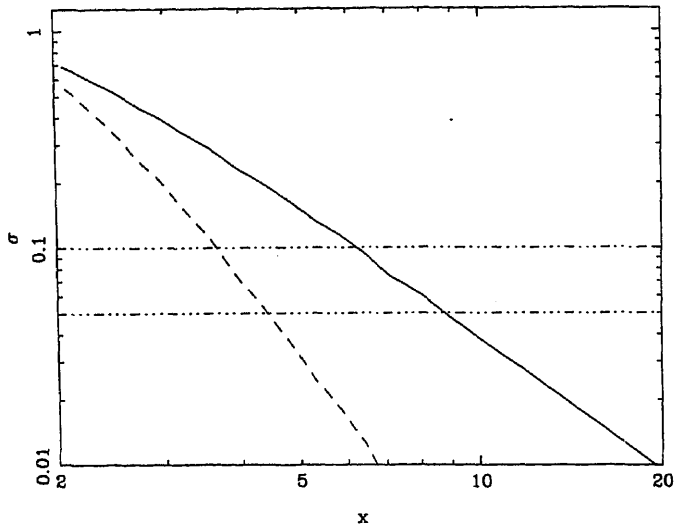


Figure 5. This figure shows rms error in numerical differentiation. We differentiated sine waves with frequencies k_x, k_y and k_z where $k_i = 2\pi/x_i$. We then compared the numerical result with its analytical counterpart to obtain an estimate of error. We have plotted the error as a function of $x = 2\pi/(k_x^2 + k_y^2 + k_z^2)^{1/2}$. Dashed line shows error for the five point formula and the thick line shows error for the mid point rule. Horizontal dot-dashed lines show the 10% and the 5% error level. This error analysis was carried out in a 64^3 box.

Table 2. This table lists the scale at which error in differentiation crosses the 5% and 10% level for the mid-point and the five point methods. A comparison of these values suggests that the five point method is more robust and less error prone of the two.

Method	5% error (k, l)	10% error (k, l)
Mid point	0.7, 9.0	1.0, 6.3
Five point	1.4, 4.5	1.7, 3.7

This translates into the following expression for a uniformly spaced grid:

$$\frac{\partial f}{\partial n} = \frac{f(n-2) - f(n+2) + 8(f(n+1) - f(n-1))}{12}. \tag{34}$$

To compare these methods we studied the variation of relative error for sine waves as a function of frequency. We computed the numerical derivatives of a set of sine waves in a 64^3 box and compared these with the analytical results. We have plotted the root mean square error as a function of the inverse of wave number $x = 2\pi/k$ in figure 5. Dashed line shows error for the five point formula and the thick line corresponds to the mid point rule. It is clear from these curves that the five point formula is far more accurate than the mid point rule.

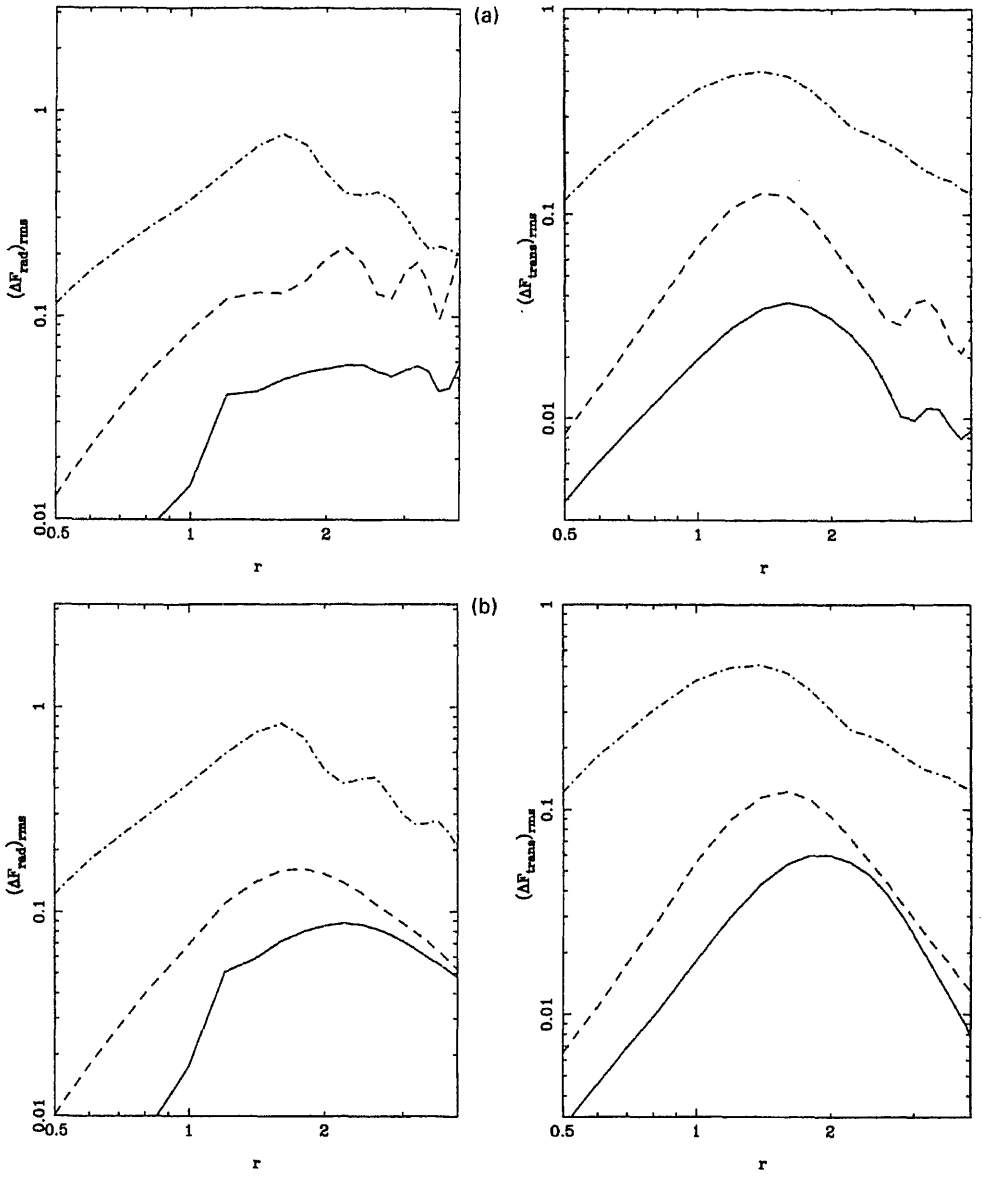


Figure 6. This figure shows the error in force of one particle. The left panel shows the root mean square dispersion in the radial force. The right panel shows corresponding curves for the transverse component of force. The thick line shows the error for TSC interpolating function, the dashed line for CIC and the dot-dashed line corresponds to NGP. In these panels the potential was computed using the poor man's Poisson solver and the gradient of the potential was computed in the Fourier space. In these panels the Poisson equation was solved using the periodic Green's function and the mid point formula was used for computing the force. Notice that the CIC and TSC curves converge in this case indicating that the error is dominated by some other component.

We have listed in table 2 the wave number at which the average error crosses some thresholds for these two formulae.

3.4 *Tests of force calculation*

In the preceding subsections we have studied the steps involved in computing the force field for a given distribution of particles. For each step we discussed alternate methods of computing the required quantity and also discussed their relative merits. We shall now discuss the errors in computing the force field for one particle. We will use many combinations of individual components like the interpolating function, Green's function and the differentiator and compare the errors for each combination.

To compute the error we placed a particle at a random position in a given lattice cell and computed force due to this at a fixed distance r in many directions to obtain the average force. We also computed the rms dispersion about the average. This process was repeated for one hundred positions of the particle in the lattice cell. Errors in this force were computed by comparing with the true $1/r^2$ force expected in such a situation. The error was split into two parts: transverse and radial. The results of this study may be summarized as follows:

- The average force deviates very little from the true force for distances larger than one grid length $[r > L]$ for almost all combinations of components.
- The root mean square dispersion about the average can be large for some combinations of components.
- Smallest rms dispersion is obtained for the combination of TSC, poor man's Poisson solver and numerical differentiation in Fourier space.
- The combination of CIC, periodic kernel and mid point differentiation also gives small errors. Surprisingly the errors in CIC and TSC (other components remaining the same) *converge* at scales larger than the grid size. This clearly indicates that the error is dominated by some other component, probably the mid-point differentiation. (Convergence disappears when five point formula or differentiation in Fourier space is used.)

To elucidate some of the points made above we have plotted the radial and transverse error for the two combinations of components mentioned in the last two points. Figure 6 shows that the error is minimum for TSC amongst the three interpolating functions in all the cases. Other points mentioned in the above discussion are also brought out clearly in these panels. We will use the combination of TSC interpolating function and the poor man's Poisson solver. We will compute the gradient of potential in the Fourier space.

4. Fast Fourier transforms

In previous sections we discussed methods used for moving the particles and the steps involved in computing the force in N-body simulations. We mentioned that fast Fourier transforms are used for computing the Fourier transforms and play a vital role in reducing

expression for the Fourier transform and reduce it to the form in which it is processed by the FFT. For details of the FFT techniques we refer the reader to any standard book on numerical analysis [15, 16].

For simplicity we will work in one dimension in this section. However we will write the d dimensional generalization of the main result. Consider a function $f(x)$ that has the Fourier transform $g(k)$. Then we may write

$$f(x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi}\, g(k)\mathrm{e}^{-ikx}. \tag{35}$$

It is not possible to use computers to integrate arbitrary functions over an infinite range in finite time. The integral is thus truncated at some finite $k_{\max}$. This truncation is possible only for those functions for which the integral over $|k| > k_{\max}$ is sufficiently small. Thus the integral to be evaluated becomes

$$f(x) \simeq \int_{-k_{\max}}^{k_{\max}} \frac{\mathrm{d}k}{2\pi}\, g(k)\mathrm{e}^{-ikx}. \tag{36}$$

This integral is evaluted in the form of a summation over the integrand

$$f(x) \simeq \sum_{k=-k_{\max}}^{k_{\max}} g(k)\mathrm{e}^{-ikx}\, \frac{\Delta k}{2\pi}, \tag{37}$$

where the function is sampled at intervals of Δk in this interval. For a constant Δk, this sampling will divide the range $-k_{\max}$ to $k_{\max}$ in N intervals where $N = (2k_{\max}/\Delta k) - 1$. The fast Fourier transforms require the number of intervals to equal an integer power of 2, which clearly is not possible for the above range if we are to sample the $k = 0$ mode as well. It is possible to circumvent this problem by making the range of integration asymmetric. We add one extra interval after $k_{\max}$ and thus make the number of intervals equal to an even number. We can further choose $k_{\max}$ and Δk to ensure that N is an integer power of two. Thus the sum may be written as

$$\begin{aligned} f(x) &\simeq \frac{1}{NL} \sum_{-k_{\max}}^{k_{\max}+\Delta k} g(k)\mathrm{e}^{-ikx} \\ &= \frac{1}{NL} \sum_{m=-N/2+1}^{N/2} g(\Delta k m)\mathrm{e}^{-i\Delta k m x}. \end{aligned} \tag{38}$$

The boxsize $NL = 2\pi/\Delta k$ defines the scale over which the function $f(x)$ is defined apart from its periodic copies. Here $N = (2k_{\max} + \Delta k)/\Delta k = 2^j$ where j is a positive integer. The function $f(x)$ is defined on a regular mesh with L as the spacing between adjacent mesh points. This allows us to rewrite the previous equation as

$$f(nL) = \frac{1}{NL} \sum_{m=-N/2+1}^{N/2} g(\Delta k m)\mathrm{e}^{-i2\pi mn/N}. \tag{39}$$

To rewrite this in the form used by the FFT algorithm we split this sum into two parts:

$$f(nL) = \frac{1}{NL} \sum_{m=-N/2+1}^{-1} g(\Delta km) e^{-i2\pi mn/N} + \frac{1}{NL} \sum_{m=0}^{N/2} g(\Delta km) e^{-i2\pi mn/N}. \quad (40)$$

Changing the summation index in the first sum by $N[m \rightarrow m+N]$ and dropping a 2π factor in the phase, we obtain

$$f(nL) = \frac{1}{NL} \sum_{m=N/2+1}^{N-1} g(\Delta k(m-N)) e^{-i2\pi mn/N} + \frac{1}{NL} \sum_{m=0}^{N/2} g(\Delta km) e^{-i2\pi mn/N}. \quad (41)$$

These terms can be put together in the following form

$$f(nL) = \frac{1}{NL} \sum_{m=0}^{N-1} G(\Delta km) e^{-i2\pi mn/N}, \quad (42)$$

where $G(\Delta km) = g(\Delta km)$ for $m \leq N/2$ and $G(\Delta km) = g(\Delta k(m-N))$ for $m > N/2$. From here it is clear that the input array for FFT should have the negative wave numbers *after* the positive wave numbers. (This can also be interpreted as 'periodicity' in k space.) The generalization of the above expression to d dimensions is

$$f(\mathbf{n}L) = \frac{1}{(NL)^d} \sum_{\mathrm{m}=0}^{N-1} G(\Delta k\mathrm{m}) e^{-i2\pi \mathrm{m.n}/N}. \quad (43)$$

Here $G(\mathrm{m})$ is the Fourier transform of the function $f(\mathbf{n}L)$ with every component $m_i > N/2$ replaced by $m_i - N$.

If an FFT routine uses a different normalization then the input array must be rescaled to get the correct amplitude for the function $f(\mathbf{n}L)$. If the FFT routine uses

$$f(\mathbf{n}L) = \frac{1}{l^d} \sum_{\mathrm{m}=0}^{N-1} G(\Delta k\mathbf{m}) e^{-i2\pi \mathrm{m.n}/N}, \quad (44)$$

then we must scale the input function as $G^{\mathrm{FFT}} = G*(l/NL)^d$. This scaling is not important if we are computing both the forward and the inverse transform as the overall normalization must be same for all FFT routines. This scaling becomes relevant only in those cases where we are transforming a given quantity only once, as in the generation of initial potential.

The above equations establish the form in which an array must be passed to the FFT routines in order to compute the Fourier transform.

5. Setting up initial conditions

5.1 *The initial density and velocity field*

N-Body simulations are generally started from fairly homogeneous initial conditions, i.e. the density contrast is much smaller than unity at all scales that can be studied using the simulation. In this regime we can use linear theory to compute all quantities of interest.

In linear theory, the evolution of density contrast can be described as a combination of a growing and a decaying mode. We can write the solution for perturbations in an Einstein–deSitter universe as

$$\delta(a) = c_1 a + c_2 a^{-3/2}. \tag{45}$$

We can evaluate c_1 and c_2 by using the initial conditions. These conditions can be written in terms of the initial velocity field and the initial potential

$$\delta_{\rm in} = a_{\rm in}\nabla^2\psi_{\rm in},$$
$$\left(\frac{\partial\delta}{\partial a}\right)_{\rm in} = -\nabla.\mathbf{u}_{\rm in}. \tag{46}$$

Using these we can express the linear solution in terms of the initial potential and the initial velocity field as

$$\delta(a) = [\tfrac{3}{5}\nabla^2\psi_{\rm in} - \tfrac{2}{5}\nabla.\mathbf{u}_{\rm in}]a + \tfrac{2}{5}[\nabla^2\psi_{\rm in} + \nabla.\mathbf{u}_{\rm in}]a^{-3/2}. \tag{47}$$

This equation suggests that for a system in growing mode, $\mathbf{u}_{\rm in} = -\nabla\psi_{\rm in}$. This equation gives us the density contrast at each point in terms of the initial gravitational potential. (We will discuss the problem of generating the initial potential in the next subsection.) Given an initial potential the above equation suggests two schemes for generating the initial density field. These are:

- The particles are distributed uniformly and their masses are proportional to the local value of density. We can either start with zero velocities, in which case we have to increase the amplitude of $\psi_{\rm in}$ by a factor $5/3$ to account for the presence of decaying mode. Alternatively we can choose to put the system in the growing mode and assign velocity $\mathbf{u}_{\rm in} = -\nabla\psi_{\rm in}$ to each particle. One major drawback of this method is that an extra array containing masses of particles needs to be stored and this can be a problem for large simulations.
- Starting with a uniform distribution the particles are displaced by a small amount, say much less than the inter-particle separation, using velocity $\mathbf{u}_{\rm in} = -\nabla\psi_{\rm in}$. The resulting distribution of particles will represent the required density field if the initial distribution did not have any inhomogeneities. We can also retain the initial velocity field.

Thus the main requirement from initial positions of particles before we generate the required perturbation is that the distribution of particles sample the potential 'uniformly.' Any inhomogeneities present in the initial distribution will combine with the density perturbations that are generated by displacing particles and will modify the initial power spectrum.

One obvious solution that ensures a uniform distribution is placing particles on the grid. This generates a distribution that is uniform but not random. An additional problem with this distribution is the extreme regularity which leads to shot noise at Nyquist frequency for potentials with large coherence length when the number of particles is smaller than the number of cells in the mesh (see figure 7).

Another 'obvious' solution is to put particles at random inside the simulation box. This distribution is uniform but it has $\sqrt{N}$ fluctuations which result in spurious clustering. The fluctuations in number of particles in a cell of size R decrease as $1/R^{3/2}$ with scale. (Comparing with power law spectra we note that this is an $n = 0$ spectrum.)

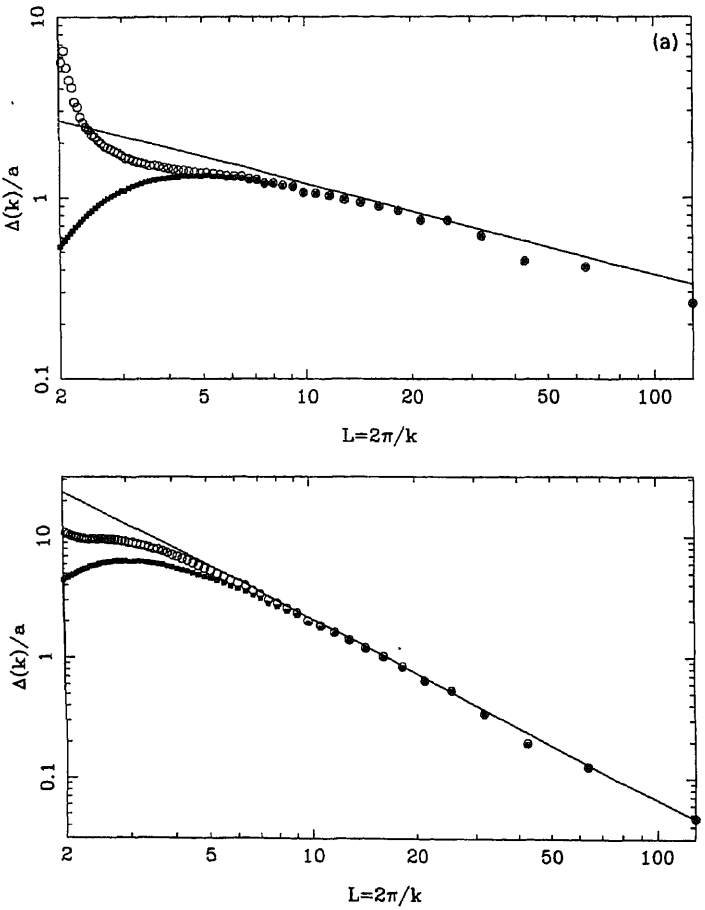


Figure 7a. This figure shows the theoretical power spectrum and the induced power spectra, linearly extrapolated to $a = 1$. The theoretical power spectrum is shown as a thick line and the power spectrum generated by displacing the particles is shown as filled squares (for simulations with 128^3 particles) and as empty circles (for simulations with 64^3 particles). The upper panel here corresponds to the power law model with $n = -2$ and lower panel to the model with $n = 0$.

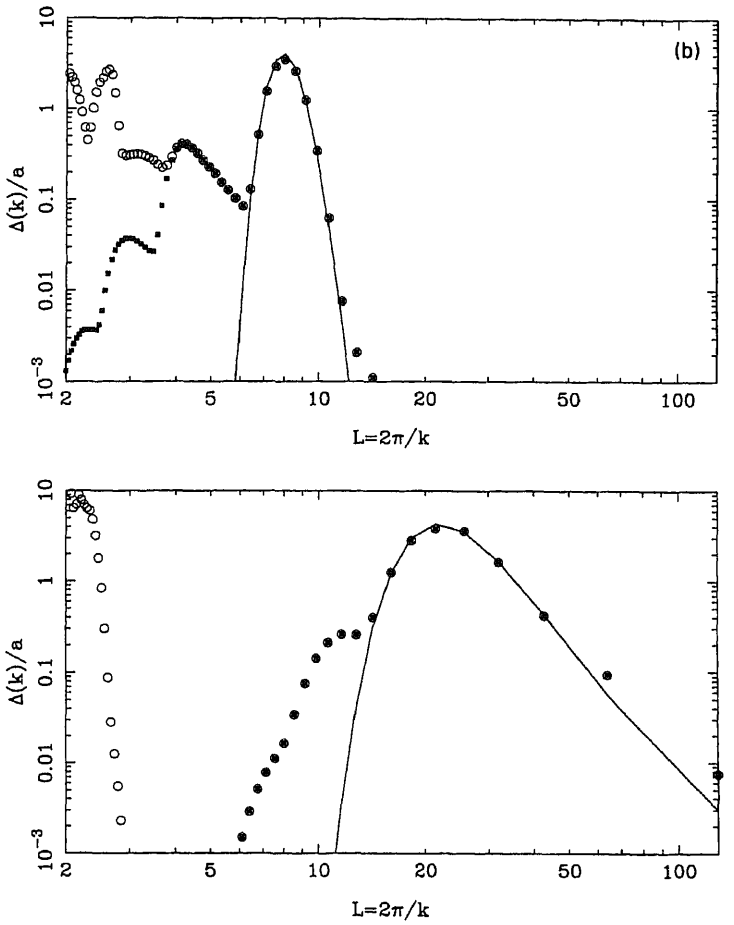


Figure 7b. These panels show the initial power spectra for models with Gaussian power spectrum. The upper panel corresponds to a Gaussian power spectrum with peak at $L_0 = 8L$ and the lower panel corresponds to a spectrum with $L_0 = 24L$. The secondary peaks at small scales occur at the harmonics of the wavenumber corresponding to the peak of the initial Gaussian.

Apart from these simple minded solutions there exist (at least) two other distributions of particles which may be used for the initial positions.

The Glass initial conditions are obtained by running a random distribution of particles through N-body with a repulsive force. In this case we start with small perturbations and the repulsive force leads to a slow decrease in the amplitude of perturbations. As the perturbations get progressively smaller it is worthwhile to estimate the evolution of inhomogeneities using linear theory. The equation for evolution of perturbations in this case is same as the standard equation for linear growth of perturbations [17] but with a source term that has an opposite sign. This leads to the following oscillatory solution with a decaying amplitude

$$\delta(a) = a^{-1/4}\left[\alpha\cos\left(\frac{\sqrt{23}}{4}\ln a\right) + \beta\sin\left(\frac{\sqrt{23}}{4}\ln a\right)\right], \tag{48}$$

where α and β are constants to be fixed by initial conditions. The process of generating initial conditions can take a long time for a large number of particles. However, this has to be done only once and the data can be stored in a file for repeated use.

The other method is a variant of the grid and Poisson initial positions mentioned above. Here particles are placed in lattice cells but at a random point within the cell. This removes the regularity of grid without sacrificing uniformity. The number fluctuation falls faster than $\sqrt{N}$ and have a smaller amplitude even at the smallest scale. The amplitude of fluctuations can be controlled by reducing the amplitude of displacement about the mesh point.

After the initial density field has been generated the initial velocities for N-body can be set as $\mathbf{u}_{\rm in} = -\nabla\psi_{\rm in}$ if we want the system to be in the growing mode. However, generation of density field from the initial potential, as outlined above involves many numerical operations that modify the potential at scales comparable to the mesh size. A better method, for consistency of density, velocity and potential, is to recompute the potential after generating the density field and set the velocities with the recomputed potential [6]. This is particularly important in case of models with a lot of small scale power. If we use the initial potential for fixing velocities then velocities have a stronger small scale component than would be expected from the density field that has been generated, leading to inconsistency in the input configuration for N-body. Therefore, we use $\mathbf{u}_{\rm in} = -\nabla\psi$, where $\nabla^2\psi = \delta(a)/a$ is the potential generated from displaced distribution of particles.

5.2 *Initial potential*

In this section we shall outline the method used to generate the initial potential. The same method can also be used to compute the density field directly if the initial density field is to be generated by placing particles with different masses on the mesh.

In most models of structure formation the initial density field is assumed to be a Gaussian random field. Linear evolution does not modify the statistics of density fields except for evolving the amplitude. As the potential and density contrast are related through a linear equation, it follows that the gravitational potential ψ is also a Gaussian random field at early epochs.

A Gaussian random field is completely described in terms of its power spectrum. The Fourier components of a Gaussian random field (both the real and the imaginary part) are random numbers with a normal distribution with variance proportional to the power spectrum of the random field. The proportionality constant depends on the Fourier transform convention. To fix this constant we consider the probability functional for a function g as

$$P[g] = B\exp\left[-\frac{1}{2}\int\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\frac{|g(k)|^2}{P_g(k)}\right]. \tag{49}$$

Here B is a normalization constant. In the Fourier tarnsfrom convention used in §3 this

equation can be written as

$$P[g] = B \exp\left[-\frac{1}{2}\frac{1}{(NL)^3}\sum_{\text{all k}}\frac{|g(k)|^2}{P_g(k)}\right]. \tag{50}$$

Thus the variance of the real and imaginary components for each Fourier mode of a Gaussian random field should equal

$$\langle |g_{\text{real}}(k)|^2\rangle = \langle |g_{\text{imag}}(k)|^2\rangle = \frac{P_g(k)N^3L^3}{2}. \tag{51}$$

The factor of 2 takes into account the fact that both the real and the imaginary components share the variance. Thus the function g may be written as

$$g(k) = (a_k + ib_k)\left[\frac{P_g(k)N^3L^3}{2}\right]^{1/2}, \tag{52}$$

where a_k and b_k are Gaussian random numbers with zero mean and unit variance.

To generate the gravitational potential we substitute the gravitational potential ψ_k in place of $g(k)$ and the power spectrum with the power spectrum for the gravitational potential. The power spectrum for the gravitational potential is $P_\psi(k,a) = P_\delta(k,a)/(a^2k^4) = P_\delta^{\text{lin}}(a=1,k)/k^4$. Here $P_\delta^{\text{lin}}(a=1,k)$ is the linearly extrapolated power spectrum of density fluctuations. With this, we can write

$$\psi_k = (a_k + ib_k)\left[\frac{P_\delta^{\text{lin}}(a=1,k)N^3L^3}{2k^4}\right]^{1/2}. \tag{53}$$

Here we have taken $a(t_0) = 1$. To specialize to the Fourier transform convention used in the FFT routine, we use $l = N^{1/2}$ (eq. (44)). This implies

$$\psi_k^{\text{FFT}} = \frac{\psi_k}{N^{3/2}L^3} = (a_k + ib_k)\left[\frac{P_\delta(k)}{2L^3k^4}\right]^{1/2}. \tag{54}$$

Here $k = 2\pi m/(NL)$. If $P_\delta(k)$ is of the form $Ak^n f(k)$, where f is a dimensionless function (it may be the transfer function or a cutoff imposed by hand) and A is the amplitude. In the units we are using here $L = 1$, therefore

$$\psi_m^{\text{FFT}} = (a_m + ib_m)\left[\frac{A}{2}\left(\frac{2\pi m}{N}\right)^{n-4} f\left(\frac{2\pi m}{N}\right)\right]^{1/2}, \tag{55}$$

where all lengths are written in units of L.

5.3 *Testing initial conditions*

In this section we discuss tests of the initial conditions generated by the method outlined above. We will compare the power spectrum of density perturbations generated by this method with the theoretical power spectrum for a wide range of models. This comparison was carried out using simulations done with 128^3 particles in a 128^3 box. To study the

effect of a smaller number of particles we also studied the power spectrum for a simulation with 64^3 particles in a 128^3 box. We use the following models for this purpose:

- Power law models with $n = -2$ and $n = 0$.
- A Gaussian power spectrum with spread $\sigma_k = 2\pi/NL$ peaked at $k_0 = 2\pi/L_0$.

$$P(k) \propto \exp\{-(k - k_0)^2/2\sigma_k^2\}.$$

This model was studied for $L_0 = 8$ and 24.

Models outlined in the first item have power at all scales whereas for the Gaussian power spectrum the power is concentrated in a narrow range of scales. This tests the accuracy of initial condition generation for the two extreme classes of models.

We have plotted the theoretical power spectrum and the induced power spectra, linearly extrapolated to $a = 1$, in figure 7 for the models listed above. The theoretical power spectrum is shown as a thick line and the power spectrum generated by displacing the particles is shown as filled squares for simulations with 128^3 particles and as empty circles for simulations with 64^3 particles.

The simulated power spectrum agrees with the theoretical power spectrum for scales larger than about $5L$ for power law models. In this range the power spectrum in simulations with different number of particles also agrees quite well.

In case of Gaussian power spectrum the simulated power spectra also contain some power at harmonics of the main peak, i.e., at scales $L_0/2, L_0/3$, etc. The Gaussian at scale L_0 is reproduced quite well. Simulations with 64^3 particles contain spurious noise around the Nyquist scales. Apart from this difference power spectra in two cases agree quite well.

Simulations with 64^3 particles generally have more noise at smaller scales. For hierarchical models this is essentially shot noise. However for models with large scale coherence (as in case of Gaussian power spectrum with $L_0 = 24L$ or $n = -2$ power law model) this can lead to large spurious noise at the Nyquist scale. This can be understood in terms of the large coherence length in the displacement field for such models, which leads to a sequence of alternately filled and empty cells. This contributes to the excess noise at the Nyquist scale.

The theoretical power spectrum is reproduced correctly in a large range of scales for a wide variety of models. Power spectrum at scales smaller than about $5L$ does not match with the theoretical power spectrum and hence results in this regime should not be used at early times. This constraint applies to all types of N-body codes as the same method is used to generate the initial conditions for all types of N-body codes. This constraint does not apply at late times as the power at small scales at late times is dictated by initial power at larger scales [18].

6. Tests of N-body code

In this section we will test performance of the N-body code as a whole. We will compare the evolution of gravitational clustering in N-body simulations with analytical results. The number of analytical results available in the non-linear regime is very limited, and hence the number of tests is also very small.

Table 3. This table lists the errors in positions and velocities for one dimensional collapse. The errors are computed by comparing with Zel'dovich approximation. The errors have been listed for three different initial perturbations, the first column contains the wave number (in units of the fundamental wave number $k_f = 2\pi/N$) of the sine wave perturbation. The errors were computed at the time of first shell crossing in the system. The N-body simulation used for this was carried out using 64^3 particles.

k/k_f	σ_x	σ_v
1	0.021	0.034
2	0.048	0.073
4	0.104	0.138

As a test of evolution of positions and velocities we can compare the motion for a one dimensional collapse before shell crossing with Zel'dovich approximation [19]. This test was suggested by Efstathiou *et al* [6] and they compared the positions and velocities for such a collapse with the corresponding quantities in Zel'dovich approximation. They chose a sine wave along one of the coordinate axes as the initial perturbation in potential. Recently [9] this test has been generalized so that the plane wave is not along any coordinate axis. In such a case the particles can approach each other with a small nonzero impact parameter. Monitoring velocities along directions perpendicular to the perturbation provides a good test of collisionless evolution. Collisionless evolution of a one dimensional perturbation cannot induce velocities along directions orthogonal to the perturbation. The authors find that particle mesh codes are the only ones that pass this test. P^3M and tree codes can also provide collisionless evolution if the softening length is of the same order as the inter particle separation.

We carried out this test with one dimensional perturbation along the x axis. The results are listed in table 3. Here we have tabulated the root mean square error in displacement and velocity. These are defined here as

$$\sigma_x = \left[\frac{\sum_i (x_i - x_i^Z)^2}{\sum_i (x_i^Z - q_i)^2}\right]^{1/2},$$

$$\sigma_v = \left[\frac{\sum_i (v_i - v_i^Z)^2}{\sum_i (v_i^Z)^2}\right]^{1/2}. \tag{56}$$

The table lists errors for three wavenumbers. In each case we have given the error at the epoch of first shell crossing in the system. The error increases with the wavenumber of the sine wave used for initial potential as we sample a given wave with a smaller number of particles and the discretization effects become more important.

The above test checks the working of an N-body code for a very special case. In a more general situation it is difficult to compare the output of numerical simulation at the level of positions of particles and only a statistical comparison is possible. We will use the averaged correlation function $\bar{\xi}$ and the scaled pair velocity h [17] for testing the accuracy of the N-body code.

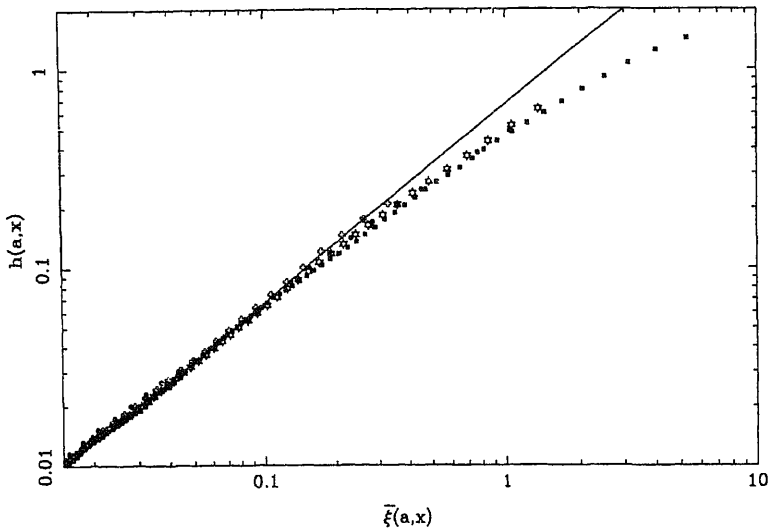


Figure 8. This figure shows $h(a,x)$ as a function of $\bar{\xi}(a,x)$ for two power law models ($n = 0$ and -1) and the CDM model. The thick line depicts the relation between these quantities in linear theory and the results from simulations are shown as points. All the points in the range $\bar{\xi} \ll 1$ lie along the line with little dispersion.

In linear regime the averaged correlation function $\bar{\xi}$ evolves as a^2, i.e., $\bar{\xi}(x,a) = (a^2/a_i^2)\bar{\xi}(x,a_i)$. In addition we know that $\bar{\xi} \ll 1$ in this regime. These can be used, along with the pair conservation equation [17] to show that $h = 2\bar{\xi}/3$ for $\bar{\xi} \ll 1$. This equation is valid for all models in linear theory. We have plotted $h(a,x)$ as a function of $\bar{\xi}(a,x)$ in figure 8. We have plotted the line corresponding to the linear theory result and have shown results from N-body simulations as points. We used simulations of power law models [$n = 0$ and -1] and the standard CDM model for this plot. Points corresponding to different simulations have been shown with different symbols. All the points crowd around the line $h = 2\bar{\xi}/3$ in the linear regime with little dispersion. The points deviate from this line for large $\bar{\xi}$ but the dispersion between different models remains small. This corresponds to an approximate 'universality' in the relation between h and $\bar{\xi}$ [20,21].

Self-similar evolution of $\bar{\xi}$ for power law models [17] can be used to test correctness of the non-linear evolution of gravitational clustering. In figure 9 we have plotted $\bar{\xi}(a,x)$ as a function of x/x_{nl} for the $n = 0$ power law model. Here x_{nl} is the scale where the linearly extrapolated $\bar{\xi}$ is unity. This scale is proportional to $a^{2/3}$ for the $n = 0$ power law model. If the evolution is self-similar then the points from different epochs must lie along a single curve. Figure 9 clearly shows that the evolution of averaged correlation function follows a self-similar pattern over a large range of scales.

The last test we consider here compares the averaged correlation function for simulations of the standard CDM model. We compare $\bar{\xi}^{1/2}$ for three simulations carried out with box size 64, 128 and 256 h^{-1} Mpc. Figure 10 shows that the curves match in the overlapping region, implying that the numerical simulation is not introducing any scale in the non-linear evolution of density perturbations. We also compare these curves with $\bar{\xi}^{1/2}$

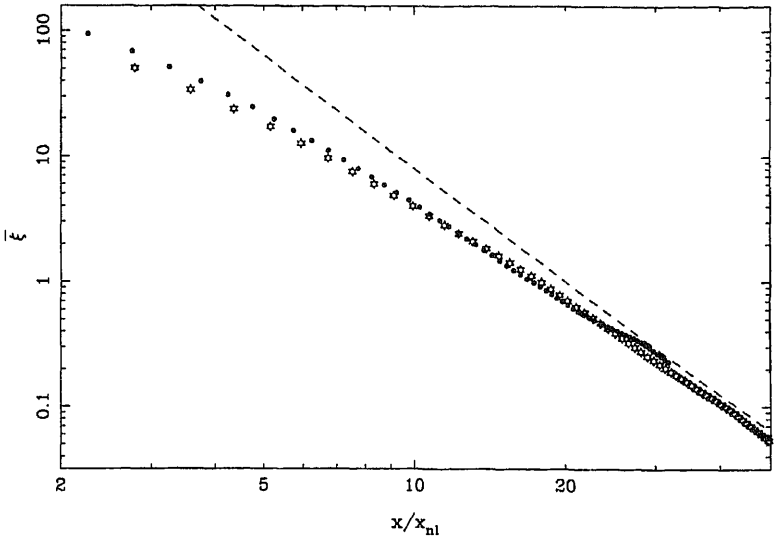


Figure 9. This figure shows $\bar{\xi}(a,x)$ as a function of x/x_{nl} for the $n=0$ power spectrum. The points have been plotted for two epochs. The overlap between points clearly shows that the system is evolving in a self similar manner. The dashed line shows the linear slope of the averaged correlation functions.

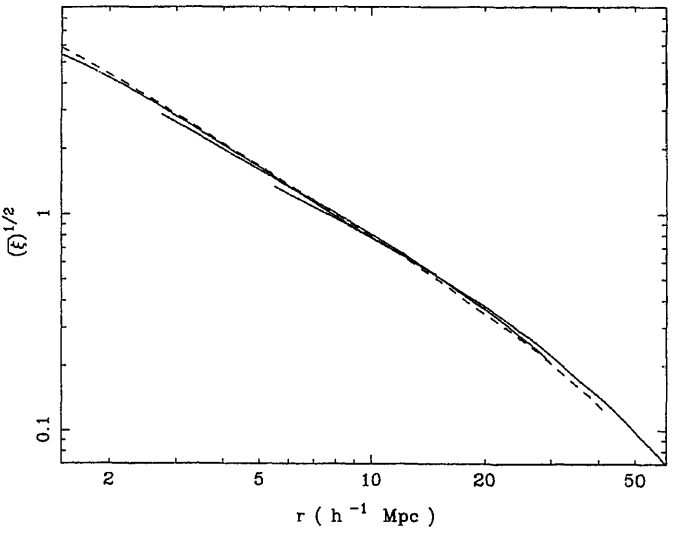


Figure 10. Comparison of simulations at different scales. This figure shows $\bar{\xi}^{1/2}$ for three CDM simulations done with box size of 64, 128 and 256 h^{-1}Mpc. Matching of curves in the overlapping region shows that the numerical simulation is not introducing any artifacts. The dashed line shows the same function from a P^3M simulation by Brieur *et al* [3]. This simulation was done with the same initial power spectrum but a different realization of the Gaussian random field was used. The similarity between $\bar{\xi}$ in the three PM simulations and the P^3M simulation shows that the non-linear evolution is being followed in the same manner in both the simulations.

obtained from a different simulation. This reference simulation was done by Brieur, Summers and Ostriker [3] using a P^3M code implemented on a GRAPE machine. This simulation was done with the same theoretical power spectrum but using a different realization of the initial density field. The two types of simulations being compared have virtually nothing in common as far as the implementation of the mathematical model is concerned. The similarity in these curves implies that both the codes are evolving density perturbations in the same manner.

7. Analysis of N-body data

In this section we will outline the method used for computing the correlation function and the scaled pair velocity [17] from simulation data.

The averaged pair velocity h is defined as

$$h(a,x) = \left\langle \frac{a(\mathrm{d}\mathbf{x}_i/\mathrm{d}a - \mathrm{d}\mathbf{x}_j/\mathrm{d}a)\cdot(\mathbf{x}_i - \mathbf{x}_j)}{(\mathbf{x}_i - \mathbf{x}_j)^2} \right\rangle_{i\neq j} = \frac{\sum_{i\leq j} \frac{a(\mathrm{d}\mathbf{x}_i/\mathrm{d}a - \mathrm{d}\mathbf{x}_j/\mathrm{d}a)\cdot(\mathbf{x}_i - \mathbf{x}_j)}{(\mathbf{x}_i - \mathbf{x}_j)^2}}{\sum_{i\leq j} 1}. \tag{57}$$

In this equation subscripts label particles and the averaging is over all pairs with separation equal to x, i.e., $x^2 = (\mathbf{x}_i - \mathbf{x}_j)^2$. This quantity is computed from simulation data by binning the pairs by pair separation x. The number of pairs in a given bin as well as the quantity to be averaged is summed for each pair in each bin. The ratio of these quantities for each bin gives the value of $h(a,x)$.

The correlation function was defined in § 1.4 as the Fourier transform of the power spectrum. This is equivalent to the following definition

$$\xi(a,x) = \langle \delta(\mathbf{y})\delta(\mathbf{z})\rangle; \quad x = |\mathbf{y} - \mathbf{z}|, \tag{58}$$

where the average is over all pairs of points with separation $\mathbf{x}$ and over all $\mathbf{x}$ with the same magnitude. This can be rewritten as

$$\xi(a,x) = \left\langle \left(\frac{\varrho(a,\mathbf{y})}{\bar{\varrho}(a)} - 1\right)\left(\frac{\varrho(a,\mathbf{z})}{\bar{\varrho}(a)} - 1\right)\right\rangle = \left\langle \left(\frac{\varrho(a,\mathbf{y})}{\bar{\varrho}(a)}\frac{\varrho(a,\mathbf{z})}{\bar{\varrho}(a)}\right) - 1\right\rangle. \tag{59}$$

Here we have used the fact that density contrast is a random field with zero mean. We can replace the densities by number densities and the product of number densities at points separated by a given distance by the number of pairs with that separation. With this the above equation reduces to

$$\xi(a,x) = \frac{n(a,x)}{\bar{n}(a)} - 1, \tag{60}$$

where $n(a,x)$ is the number of pairs with separation x and $\bar{n}$ is the number of pairs in a uniform distribution. Therefore, the problem is again reduced to computing the number of

pairs separated by a given distance. This is done by dividing the range of x into small intervals and binning the pairs into these intervals. Similar operation is required for computing the scaled pair velocity, therefore the binning operation for computing these quantities can be combined conveniently.

We can obtain an expression for the averaged correlation by averaging ξ.

$$\bar{\xi}(a,x) = \frac{3}{x^3}\int^x \mathrm{d}y y^2 \xi(a,y) = \frac{\int^x \mathrm{d}y y^2 \xi(a,y)}{\int^x \mathrm{d}y y^2}$$
$$= \frac{\int^x \mathrm{d}y n(y)}{\int^x \mathrm{d}y \bar{n}(y)} - 1, \tag{61}$$

where we have used the fact that the average number of pairs with a given separation is proportional to the square of the separation for a uniform distribution of particles. In the discrete realization of this expression the integrals over number of pairs are replaced by summation over bins. Computationally this quantity is less error prone than the correlation function. As for large separations, the number of pairs in a given bin in the simulation output approaches the number of pairs for a uniform distribution. Subtracting two nearly equal numbers can give large error. However, in case of the averaged correlation function the excess number of pairs at small scales is carried over to large scales through summation over all smaller scales. And in general the ratio $(\sum n)/(\sum \bar{n})$ is larger at all scales in comparison with $n/\bar{n}$.

8. Discussion

In the preceeding sections we have described the basic mathematical model that is implemented in particle mesh cosmological N-body codes. We have not discussed the P^3M codes, tree codes, etc. in this review as only PM codes are known to ensure collisionless evolution [9] though they suffer from a very limited resolution. Other methods improve spatial resolution but do not ensure collisionless evolution. However, most of the machinery is common to these codes and many of the results can be carried over to these codes. Anyone interested in using our N-body code may contact the authors.

Acknowledgements

The authors thank S F Shandarin and F R Bouchet for useful discussions on many aspects of N-body codes. JSB thanks CSIR, India for the Senior Research Fellowship.

References

[1] R W Hockney and J W Eastwood, *Computer Simulations using Particles* (IOP Publ., 1988)
[2] F R Bouchet and L Hernquist, *Astrophys. J. Sci.* **68**, 521 (1988)
[3] P P Brieu, F J Summers and J P Ostriker, *Astrophys. J.* **433**, 566 (1995)
[4] F R Bouchet, J C Adams and R Pellat, *Astron. Astrophys.* **144**, 413 (1984)
[5] F R Bouchet and H Kandrup, *Astrophys. J.* **299**, 1 (1985)
[6] G Efstathion, M Davis, C S Frenk and S D M White, *Astrophys. J. Sci.* **57**, 241 (1985)

[7] J Barnes and P Hut, *Nature (London)* **324**, 446 (1986)
[8] L Hernquist, F R Bouchet and Y Suto, *Astrophys. J. Sci.* **75**, 231 (1991)
[9] A L Mclott, S F Shandarin, R J Splinter and Y Suto, *Astrophys. J.* (in Press) (1997)
[10] I Suisalu and E Saar, *Mon. Not. R. Astron. Soc.* (1996) (submitted)
[11] A A Klypin and S F Shandarin, *Mon. Not. R. Astron. Soc.* **204**, 891 (1983)
[12] W M Irvine, *Doctoral Dissertation* (Harvard University, 1961)
[13] D Layzer, *Astrophys. J.* **138**, 174 (1963)
[14] N A Dmitriev and Ya B Zel'dovich, *Sov. Phys. JETP* **18**, 793 (1964)
[15] H M Antia, *Numerical methods for scientists and engineers* (Tata McGraw Hill, 1991)
[16] W H Press, B P Flannery, S A Teukolsky and W T Vetterling, *Numerical Recipes* (Cambridge University Press, 1986)
[17] P J E Peebles, *Large scale structure of the universe* (Princeton University Press, 1980)
[18] J S Bagla and T Padmanabhan, 1996, astro-ph/9605202, IUCAA preprint 20/96, submitted to *Mon. Not. R. Astron. Soc.*
[19] Ya B Zel'dovich, *Astron. Astrophys.* **5**, 84 (1970)
[20] A J S Hamilton, P Kumar, Lu Edward and A Matthews, *Astrophys. J.* **374**, L1 (1991)
[21] R Nityananda and T Padmanabhan, *Mon. Not. R. Astron. Soc.* **271**, 976 (1994)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 2,
August 1997
pp. 193–197

Differentiability of invariants and Liapunov equations in Hamiltonian systems

B R SITARAM
Physical Research Laboratory, Navrangpura, Ahmedabad 380009, India

MS received 27 April 1995; revised 7 April 1997

Abstract. In this communication, we investigate the behavior of the derivatives of invariants for Hamiltonian systems, using information derived from an analysis of the Liapunov exponents of the system. We show that under certain conditions on the analyticity properties of the solutions of the equations of motion, it is possible to construct $2n$ invariants of motion which are guaranteed to be C^∞ as functions of phase space and time in a suitably defined domain D.

Keywords. Integrability; Liapunov exponents; chaos; Hamiltonian systems.

PACS Nos 3.20; 2.40

In a previous paper [1] (referred to as I in what follows), we had shown that an analysis of the solutions of the Liapunov equations for a Hamiltonian system can give important qualitative information regarding the differentiability of invariants. In particular, we had shown that solutions of the Liapunov equations can be used to construct invariants; the value of the Liapunov exponent would then specify how the first order phase space derivatives of the invariant behave. In this paper, we continue our analysis of the problem and show how explicit equations can be written down for higher order derivatives of the invariants. It will turn out that these equations can be solved, in terms of the solutions of the Liapunov equations themselves.

Consider a Hamiltonian system with n degrees of freedom, with phase variables denoted by $z_i, i = 1, \ldots, 2n$ and Hamiltonian H. We denote the canonical Poisson bracket matrix by ω, so that the equations of motion (EOM) can be written as

$$\frac{\mathrm{d}z_i}{\mathrm{d}t} = \omega_{ij}\frac{\partial H}{\partial z_j} \equiv \omega_{ij} H_j. \tag{1}$$

The equations of Liapunov (EOL), which are the first variational equations of the EOM then read:

$$\frac{\mathrm{d}\eta_i}{\mathrm{d}t} = \omega_{ij} H_{jk}\,\eta_k, \tag{2}$$

where, in this equation and in the following subscripts on H denote derivatives with regard to the phase space variables, z_i. As we have argued in I, these equations can be reinterpreted as follows: consider η_i as functions on phase space and time; introduce the

Then, the EOL is equivalent to the PDE

$$Z\eta_i = \omega_{ij} H_{jk} \eta_k. \tag{4}$$

There exist $2n$ independent solutions of the EOL. Let $\eta_{i\mu}$ denote the $2n \times 2n$ matrix of solutions obtained in this way; we assume that these solutions have been selected so that $\det(\eta_{i\mu}) = 1$. We note the following properties of $\eta_{i\mu}$:

1. If η is a symplectic matrix, $\eta^T \cdot \Omega \cdot \eta = \Omega$, at $t = 0$, it is so at all times.
2. Given a solution matrix $\eta_{i\mu}$, we can construct invariants $F_\mu(z_i, t)$ by the equations

$$\frac{\partial F_\mu}{\partial z_i} = \Omega_{ij} \eta_{j\mu}, \tag{5}$$

$$\frac{\partial F_\mu}{\partial t} = \eta_{i\mu} H_i, \tag{6}$$

provided, at $t = 0$, we have

$$\Omega_{ik} \frac{\partial \eta_{k\mu}}{\partial z_j} = \Omega_{jk} \frac{\partial \eta_{k\mu}}{\partial z_i}. \tag{7}$$

Let $\bar{\eta}_{\mu i}$ denote the inverse matrix. If η is symplectic, then it can be easily seen that $\bar{\eta} = \omega \cdot \eta^T \cdot \Omega$. One special choice of initial conditions on η is to assume that $\eta_{i\mu}(t = 0) = \delta_{i\mu}$; for simplicity of notation, we shall assume that this choice has been made. (In the more general case, the solutions of the inhomogeneous equations given below will have to have an extra term corresponding to the initial values; our choice sets these initial values to 0.) This choice ensures that η is symplectic and that invariants can be constructed.

Given such a choice of $\eta_{i\mu}$, define first order differential operators X_μ by $X_\mu = \eta_{i\mu}(\partial/\partial z_i)$. It is then straightforward to see that $[Z, X_\mu] = 0$. This property of X_μ is crucial to our analysis, as it allows us to immediately write down the equations determining $\eta_{i\mu}$ and its derivatives along the X_μ directions:

$$Z\eta_{i\mu} = \omega_{ij} H_{jk} \eta_{k\mu}, \tag{8}$$

$$ZX_\nu \eta_{i\mu} = \omega_{ij} H_{jk} X_\nu \eta_{k\mu} + \eta_{k\mu} X_\nu \omega_{ij} H_{jk}. \tag{9}$$

We now look at these two equations once again along the characteristics, which is essentially equivalent to selecting a specific orbit of H, replacing all phase space variables by explicit functions of t and replacing Z by the operator $\mathrm{d}/\mathrm{d}t$. The second equation then is an inhomogeneous version of the first equation; also, because of our choice of initial conditions, the second equation has to be solved with the initial condition $X_\nu \eta_{i\mu} = 0$. This yields the solution for the second equation:

$$X_\nu \eta_{i\mu} = \eta_{ir} \int_0^t \bar{\eta}_{rj} \omega_{jm} H_{klm} \eta_{k\mu}. \tag{10}$$

An immediate consequence of this solution is that $[X_\mu, X_\nu] = 0$, i.e., the $2n+1$ vectors Z, X_μ form an involutive distribution. The Frobenius theorem [2] then assures us that there exists a local coordinate system, defined by $\{\tau, x^\mu, \mu = 1, \ldots, 2n\}$ such that $Z = \partial/\partial\tau, X_\mu = \partial/\partial x^\mu$. τ is the arc parameter of the orbit, while the x_μ define coordinates transverse to the orbit. (Not surprisingly, the x_μ are just a relabeling of the invariants F_μ.)

Higher order derivatives can be calculated in a similar fashion. In fact, all such derivatives are governed by inhomogeneous PDEs. Along the charateristics, the homogeneous part of the resulting ODE coincides with the EOL and we have the equations

$$ZX_{\nu_n} \cdots X_{\nu_2} X_{\nu_1} \eta_{i\mu} = \omega_{ij} H_{jk} X_{\nu_n} \cdots X_{\nu_2} X_{\nu_1} \eta_{k\mu} + R_{i\mu\nu_n \cdots \nu_2 \nu_1}, \tag{11}$$

where, we have the recurrence relationship:

$$R_{i\mu\nu_{n+1}\nu_n \cdots \nu_2\nu_1} = X_{\nu_{n+1}} R_{i\mu\nu_n \cdots \nu_2\nu_1} + (X_{\nu_n} \cdots X_{\nu_2} X_{\nu_1} \eta_{k\mu})\omega_{il} X_{\nu_{n+1}} H_{lk}. \tag{12}$$

Thus, we have the solution

$$X_{\nu_n} \cdots X_{\nu_2} X_{\nu_1} \eta_{i\mu} = \eta_{i\tau} \int_0^t \bar{\eta}_{\tau j} R_{j\mu\nu_n \cdots \nu_2\nu_1}. \tag{13}$$

We can use these equations to prove the following:

Theorem. *Assume that H is a polynomial in phase space variables. Assume also that for a given orbit, the solutions of the EOM are analytic in a neighborhood D of the real t axis. Let F_μ be the invariants constructed using the above procedure. Then these invariants are analytic (as functions of the arc parameter, t) along the orbit and are C^∞ in directions transverse to the orbit in the region D.*

Proof. The proof follows from the following steps:

1. By assumption, the solution of the EOM is analytic in some neighborhood D of the real t axis.
2. The $\eta_{i\mu}$ are solutions of linear differential equations, whose coefficients are polynomial functions on phase space, as the Hamiltonian is itself a polynomial. Thus, every real t forms an ordinary point, and hence the solutions $\eta_{i\mu}$ themselves are analytic in the same neighborhood D of the real t axis [3].
3. By induction, the inhomogeneous terms $R_{i\mu\nu_{n+1}\nu_n \cdots \nu_2\nu_1}$ are themselves analytic in the same neighborhood D of the real t axis; the $n+1$th order phase space derivatives, which are obtained by integrating such functions are thus also analytic in D.

Since the X_μ, Z form a basis for the tangent space of the extended phase space, the theorem follows.

It is simple to find examples to show that the C^∞ property of the derivatives of the $\eta_{i\mu}$ cannot be replaced by C^ω. In fact, consider the 1 d.o.f. Hamiltonian, $H = p^2/2 + q^4/4 - q^2/2$, which has an unstable equilbrium point at $q = 0, p = 0$. The corresponding solutions of the Liapunov equations are:

$$\eta_{11} = \cosh(t), \eta_{12} = \sinh(t), \eta_{21} = \sinh(t), \eta_{22} = \cosh(t). \tag{14}$$

It is possible to solve the equations (11) for the derivatives quite easily: (Note that all the odd ordered derivatives vanish)

$$X_1^2\eta_{11} = \frac{-3}{4e^t} + \frac{3e^t}{2} - \frac{3e^{3t}}{4}, \tag{15}$$

$$X_1^4\eta_{11} = \frac{15}{4e^t} + \frac{45e^t}{8} - \frac{45e^{3t}}{4} + \frac{15e^{5t}}{8} + \frac{45e^t t}{2}, \tag{16}$$

$$X_1^6\eta_{11} = \frac{5355}{32e^t} - \frac{4095e^t}{16} + \frac{1575e^{5t}}{16} - \frac{315e^{7t}}{32} - \frac{945t}{8e^t} + \frac{945e^t t}{4} - \frac{2835e^{3t}t}{8}. \tag{17}$$

Higher order derivatives have behaviors which can be inferred from the above; in fact,

$$X_1^n\eta_{11} \simeq \mathrm{e}^{(n+1)t} \tag{18}$$

whenever n is even. It is clear from the form of these derivatives that the Taylor series formed by using these derivatives has a radius of convergence which rapidly approaches 0 as $t \to \infty$.

The above counter example raises the question of whether the behavior seen above is because of the presence of exponential functions in the definition of the X_μ. For this example, it is easy to show that even the derivatives evaluated along the q,p directions show the same exponential behavior, with the exponent increasing with the order of the derivatives. We give below the results for a few of the derivatives of η_{11} along the q,p directions:

$$\frac{\partial^2}{\partial p^2}\eta_{11} = \frac{-4}{5e^{4t}} + \frac{4}{5e^t} + \frac{4e^t}{5} - \frac{4e^{4t}}{5}, \tag{19}$$

$$\frac{\partial^2}{\partial p\partial q}\eta_{11} = \frac{-4}{5e^{4t}} + \frac{4}{5e^t} + \frac{4e^t}{5} - \frac{4e^{4t}}{5}, \tag{20}$$

$$\frac{\partial^2}{\partial q^2}\eta_{11} = \frac{-4}{5e^{4t}} + \frac{4}{5e^t} + \frac{4e^t}{5} - \frac{4e^{4t}}{5}, \tag{21}$$

$$\frac{\partial^4}{\partial p^4}\eta_{11} = \frac{472}{425e^{16t}} + \frac{8}{25e^{4t}} - \frac{608}{425e^t} + \frac{8e^{4t}}{25} + \frac{472e^{16t}}{425}, \tag{22}$$

$$\frac{\partial^4}{\partial p^3\partial q}\eta_{11} = \frac{472}{425e^{16t}} - \frac{172}{25e^{4t}} + \frac{2452}{425e^t} + \frac{2452e^t}{425} - \frac{172e^{4t}}{25} + \frac{472e^{16t}}{425}, \tag{23}$$

$$\frac{\partial^4}{\partial p^2\partial q^2}\eta_{11} = \frac{472}{425e^{16t}} - \frac{232}{25e^{4t}} + \frac{3472}{425e^t} + \frac{3472e^t}{425} - \frac{232e^{4t}}{25} + \frac{472e^{16t}}{425}, \tag{24}$$

$$\frac{\partial^4}{\partial p\partial q^3}\eta_{11} = \frac{472}{425e^{16t}} - \frac{232}{25e^{4t}} + \frac{3472}{425e^t} + \frac{3472e^t}{425} - \frac{232e^{4t}}{25} + \frac{472e^{16t}}{425}, \tag{25}$$

$$\frac{\partial^4}{\partial q^4}\eta_{11} = \frac{472}{425e^{16t}} - \frac{232}{25e^{4t}} + \frac{3472}{425e^t} + \frac{3472e^t}{425} - \frac{232e^{4t}}{25} + \frac{472e^{16t}}{425}. \tag{26}$$

The peculiar behavior shown above seems at first sight puzzling, as the system under consideration is integrable. However, the Hamiltonian has zero gradient at the equilibrium point; as a result the usual result that the existence of a conserved quantity implies the vanishing of a pair of Liapunov exponents is no longer valid. Also, it is easy

to see that there is a change of topology of the solutions at the unstable equilibrium points, with a single solution bifurcating into two solutions. It is conjectured that such changes are related to the behavior encountered above. More detailed investigations of this point are in progress.

The fact that $X_1^n \eta_{11} \simeq e^{(n+1)t}$ is in fact characteristic of many systems. In fact, assume that we have a Hamiltonian H, with a periodic solution for some choice of initial conditions. By Floquet theory [3], the $\eta_{i\mu}$ are of the form:

$$\eta_{i\mu} = e^{\lambda_\mu t} \phi_{i\mu}(t), \tag{27}$$

where, λ_μ is the Floquet exponent and $\phi_{i\mu}(t)$ are periodic functions in t. Assume also that $\lambda \equiv \lambda_1 > 0$. An easy induction then shows that $X_1^n \eta_{i1}$ is dominated by terms of the form $e^{(n+1)\lambda t} \phi_i^{(n)}(t)$, where $\phi_i^{(n)}(t)$ is periodic. Examples of such unstable periodic orbits can be easily found, e.g., the special solutions corresponding to $y = 0$, $py = 0$ for the Hamiltonian $H = p_x^2/2 + p_y^2/2 + x^4/4 + y^4/4 + \alpha x^2 y^2/2$ with $\alpha < 0$.

A final remark should be made regarding the assumption that the solutions of the equations of motion should be analytic in some strip which includes the real t axis. The work of Chang *et al* [4] shows that chaotic trajectories may be characterized by singularities approaching arbitrarily close to the real t-axis. In such a case, the above proof breaks down and it is not very clear how the derivatives would behave. Attempts are in progress to address this problem numerically.

References

[1] B R Sitaram, *Pramana – J. Phys.* (1995)

[2] N J Hicks, *Notes on differential geometry* (Van Nostrand East West Press, New Jersey, 1969)

[3] E T Whittaker and G N Watson, *A course of modern analysis* (Cambridge University Press, Cambridge, 1969)

[4] Y F Chang, M Tabor and J Weiss, *J. Math. Phys.* **23**, 531 (1982)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 2,
August 1997
pp. 199–204

Truncated harmonic oscillator and parasupersymmetric quantum mechanics

B BAGCHI, S N BISWAS[1], AVINASH KHARE[2] and P K ROY[3]
Department of Applied Mathematics, University of Calcutta, 92 APC Road, Calcutta 700 009, India
[1] Department of Physics, Delhi University, Delhi 110 007, India
[2] Institute of Physics, Sachivalaya Marg, Bhubaneswar 751 005, India
[3] Department of Physics, Haldia Government College, Haldia 721 657, India

MS received 3 June 1996

Abstract. We discuss in detail the parasupersymmetric quantum mechanics of arbitrary order where the parasupersymmetry is between the normal bosons and those corresponding to the truncated harmonic oscillator. We show that even though the parasusy algebra is different from that of the usual parasusy quantum mechanics, still the consequences of the two are identical. We further show that the parasupersymmetric quantum mechanics of arbitrary order p can also be rewritten in terms of p supercharges (i.e. all of which obey $Q_i^2 = 0$). However, the Hamiltonian cannot be expressed in a simple form in terms of the p supercharges except in a special case. A model of conformal parasupersymmetry is also discussed and it is shown that in this case, the p supercharges, the p conformal supercharges along with Hamiltonian H, conformal generator K and dilatation generator D form a closed algebra.

Keywords. Harmonic oscillator; parasupersymmetry; conformal generator; dilatation generator.

PACS Nos 11.30; 03.65

A great deal of attention is now being paid to study [1–4] quantum mechanics in a finite dimensional Hilbert space (FHS). In particular, we would like to mention the recent developments [2, 3, 4] in quantum phase theory which deals [2] with a quantized harmonic oscillator in a FHS and which finds interesting applications [4] in problems of quantum optics.

Recently we studied [5] some basic properties of these oscillators. In particular it was pointed out that the raising and lowering operators of the truncated oscillator behave like parafermi oscillator. Inspired by this similarity, a parasupersymmetric quantum mechanics (PSQM) of order 2 was also written down where the parasusy is between the usual bosons and the truncated bosons. However, the explicit form of the charge was not written down. Further, the consequences were also not elaborated upon. The purpose of this note is to generalize this construction to arbitrary order. In particular, we show that for these PSQM models of arbitrary order p, the algebra is given by

$$Q^{p+1} = 0; \quad [H, Q] = 0, \tag{1}$$

$$Q^p Q^+ + Q^{p-1} Q^+ Q + \cdots + Q^+ Q^p = p(p+1) Q^{p-1} H \tag{2}$$

and the Hermitian conjugated relations and discuss their consequences in some detail. In particular, we show that the consequences following from this algebra are identical to those following from the well-known PSQM model of the same order p [6, 7] even though the two algebras are different. In particular, whereas eq. (1) is identical in the two schemes, eq. (2) is different in the two schemes in the sense that in the well-known case, the coefficient on the r.h.s. is $2p$ instead of $p(p+1)$ in eq. (2). In view of the identical consequences, it is worth examining as to why the PSQM of order p can be written down in an alternative way. To that end we show that one can in fact express PSQM of order p in terms of p super (rather than parasuper) charges all of which satisfy $Q_i^2 = 0$ and further all of them commute with the Hamiltonian. However, unlike the usual supersymmetric (SUSY) quantum mechanics (QM), here H cannot be simply expressed in terms of the p supercharges except in a special case. In the special case we show that the Hamiltonian has a very simple expression in terms of the p supercharges

$$Q_1 Q_1^+ + \sum_{j=1}^{p} Q_j^+ Q_j = 2H. \tag{3}$$

We also discuss a parasuperconformal model of order p and show that the dilatation and conformal operators also can be similarly expressed in quadratic form in terms of the p SUSY and p parasuperconformal charges.

Let us start with the truncated raising and lowering operators a^+ and a. It is well-known that if one truncates at $(p+1)$th level ($p > 0$ is an integer) then a and a^+ can be represented by $(p+1) \times (p+1)$ matrices and they satisfy the commutation relation [8]

$$[a, a^+] = I - (p+1)K, \tag{4}$$

where I is $(p+1) \times (p+1)$ unit matrix while $K = \mathrm{diag}(0, 0, \ldots, 0, 1)$ with $Ka = 0$ and further $K^2 = K \neq 0$. As shown by Kleeman [9], the irreducible representations of eq. (4) are the same as those for the scheme

$$[a, a^+ a] = a;\ a^{p+1} = 0; \quad a^j \neq 0 \quad \text{if} \quad j < (p+1). \tag{5}$$

A convenient set of representation of the matrices a and a^+ is given by

$$(a)_{\alpha,\beta} = \sqrt{\alpha}\,\delta_{\alpha+1,\beta} \tag{6}$$

$$(a^+)_{\alpha,\beta} = \sqrt{\beta}\,\delta_{\alpha,\beta+1} \tag{7}$$

where $\alpha, \beta = 1, 2, \ldots, (p+1)$. As shown in [5], the nontrivial multilinear relation between a and a^+ is given by

$$a^p a^+ + a^{p-1} a^+ a + \cdots + a a^+ a^{p-1} + a^+ a^p = \frac{p(p+1)}{2} a^{p-1}. \tag{8}$$

These relations are strikingly similar to those of parafermi oscillator of order p [7] except that in the latter case, the coefficient on the right hand side is $p(p+1)(p+2)/6$ unlike $p(p+1)/2$ in eq. (8). As expected, for the case of the Fermi oscillator ($p = 1$), both the coefficients are same while they are different otherwise.

Motivated by the nontrivial relation between a and a^+ as given by eq. (8) it is worth enquiring if one can construct a kind of PSQM of order p in which there will be symmetry

between bosons and truncated bosons of order p. It turns out that the answer to the question is yes. In particular, on choosing the parasusy charges Q and Q^+ as $(p+1)\times(p+1)$ matrices as given by

$$(Q)_{\alpha\beta} = b^+ a = \sqrt{\alpha}(P + iW_\alpha)\delta_{\alpha+1,\beta}, \tag{9}$$

$$(Q^\dagger)_{\alpha\beta} = ba^+ = \sqrt{\beta}(P - iW_\beta)\delta_{\alpha,\beta+1}, \tag{10}$$

where b, b^+ denote the bosonic annihilation and creation operators and $\alpha, \beta = 1, 2, \ldots, (p+1)$, so that Q and $Q^\dagger$ automatically satisfy $Q^{p+1} = 0 = (Q^\dagger)^{p+1}$. Further, it is easily shown that the Hamiltonian $\hbar = m = 1$)

$$(H)_{\alpha\beta} = H_\alpha \delta_{\alpha\beta}, \tag{11}$$

where $(r = 1, 2, \ldots, p)$

$$H_r = \frac{P^2}{2} + \frac{1}{2}(W_r^2 - W_r') + \frac{1}{2}C_r,$$

$$H_{p+1} = \frac{P^2}{2} + \frac{1}{2}(W_p^2 + W_p') + \frac{1}{2}C_p \tag{12}$$

commutes with the parasusy charges Q and $Q^\dagger$ (i.e. $[H, Q] = 0 = [H, Q^+]$) provided $(s = 2, 3, \ldots, p)$

$$W_{s-1}^2 + W_{s-1}' + C_{s-1} = W_s^2 - W_s' + C_s. \tag{13}$$

Here $C_1, C_2, \ldots, C_p$ are arbitrary constants with the dimension of energy. It turns out that the nontrivial relation given by eq. (2) between $Q, Q^\dagger$ and H is satisfied provided

$$C_1 + 2C_2 + \cdots + pC_p = 0. \tag{14}$$

It is interesting to notice that the parasusy charge as well as the algebra as given by eqs (1), (2), (9), (10) and (14) is very similar to that of standard PSQM of order p [7] except that in the standard case the coefficient on the r.h.s. of eq. (2) is $2p$ instead of $p(p+1)/2$ and instead of eq. (14) in the standard case one has

$$C_1 + C_2 + \cdots + C_p = 0. \tag{15}$$

Besides, unlike in eq. (9), in the standard case, Q is defined without the factor of $\sqrt{\alpha}$. However, the Hamiltonian and the relation between the superpotentials as given by eqs (11) to (13) are identical in the two cases. As a result the consequences following from the two different PSQM schemes of order p are identical. In particular, as shown in [7], in both the cases (i) the spectrum is not necessarily positive semidefinite unlike in SUSY QM, (ii) the spectrum is $(p+1)$-fold degenerate atleast above the first p levels while the ground state could be $1, 2, \ldots, p$ fold degenerate depending on the form of the superpotentials and (iii) one can associate p ordinary SUSY QM Hamiltonians.

Why do the two seemingly different PSQM schemes given the same consequences? The point is that in the case of parasusy of order p, one has p independent parasusy charges and in the two schemes one has merely used two of the p independent forms of Q. It is then clear that one can in fact define p seemingly different PSQM schemes of order p but all of them will have identical consequences. For example, the parafermi operators

are usually defined by the following $(p+1)\times(p+1)$ matrices [7]

$$(a)_{\alpha\beta} = \sqrt{\alpha(p-\alpha+1)}\delta_{\alpha+1,\beta}, \tag{16}$$

$$(a^+)_{\alpha,\beta} = \sqrt{\beta(p-\beta+1)}\delta_{\alpha,\beta+1}. \tag{17}$$

So one could as well have defined the parasusy charges by

$$(Q)_{\alpha\beta} = b^+a = \sqrt{\alpha(p-\alpha+1)}(P+iW_\alpha)\delta_{\alpha+1,\beta}, \tag{18}$$

$$(Q^\dagger)_{\alpha\beta} = ba^+ = \sqrt{\beta(p-\beta+1)}(P-iW_\beta)\delta_{\alpha,\beta+1} \tag{19}$$

instead of the usual choice without the square root factor [7]. It is easily shown that in this case too the parasusy charges Q and Q^+ satisfy the algebra as given by eqs (1), (2) and (14) except that the factor on the r.h.s. of eq. (2) is now $p(p+1)(p+2)/3$ and the constants C_i satisfy

$$p(C_1+C_p)+2(p-1)(C_2+C_{p-1})+\cdots + \frac{(p+1)}{2}C_{(p+1)/2}=0, \quad p \text{ odd}, \tag{20}$$

$$p(C_1+C_p)+2(p-1)(C_2+C_{p-1})+\cdots + \frac{p(p+2)}{4}C_{p/2}+C_{(p+2)/2}=0, \quad p \text{ even} \tag{21}$$

instead of eq. (14). However, as before the Hamiltonian and the relation between the various superpotentials is unaltered and hence one would get the same consequences as in the standard PSQM case [7].

At this stage, it is worth asking if parasusy QM of order p can be put in an alternative form by making use of the fact that there are p independent parasupercharges [7]? If yes this would be analogous to the so called Green construction for parafermi and parabose operators [10]. We now show that the answer to the question is yes. Let us first note that the supercharge as given by eq. (9) can be written down as a linear combination of the following p supercharges [12]

$$Q = \sum_{j=1}^{p}\sqrt{j}Q_j, \tag{22}$$

where

$$(Q_j)_{\alpha\beta} = (P-iW_j)\delta_{\alpha+1,\beta=j+1}. \tag{23}$$

It is easily checked that these p charges Q_j are in fact supercharges in the sense that all of them satisfy $Q_j^2=0$. Further, all of them commute with the Hamiltonian as given by eq. (11) provided condition (13) is satisfied. Besides they satisfy

$$Q_iQ_j = 0 \quad \text{if} \quad j\neq i+1, \tag{24}$$

$$Q_iQ_j^+ = 0 = Q_i^+Q_j \quad \text{if} \quad i\neq j. \tag{25}$$

However, there is one respect in which these charges are different from the usual SUSY charges: the nontrivial relation of the usual parasusy algebra (i.e. eq. (2) but with $2p$ on

the r.h.s. instead of $p(p+1)$) now contains product of all p charges i.e.

$$Q_1[Q_1^+Q_1Q_2\cdots Q_{p-1}+Q_2Q_2^+Q_2\cdots Q_{p-1}+\cdots+Q_2Q_3\cdots Q_pQ_p^+]$$
$$=2pQ_1Q_2\cdots Q_{p-1}H, \tag{26}$$
$$[Q_1^+Q_1Q_2\cdots Q_{p-1}+Q_2Q_2^+Q_2\cdots Q_{p-1}+\cdots+Q_2Q_3\cdots Q_pQ_p^+]Q_p$$
$$=2pQ_2Q_3\cdots Q_pH, \tag{27}$$

provided eq. (15) is satisfied. If one considers other versions of PSQM of order p, then one would have similar relations but with different weight factors between the various terms and also different relations between C_i which can easily be worked out.

There is one special case however when the algebra takes a particularly simple form. In particular when all the constants C_i are zero then it is easily checked that the Hamiltonian can be written as a sum over quadratic pieces in Q as given by eq. (3) which is a generalization of the SUSY algebra in the case of p supercharges. In this case, clearly the spectrum is positive semidefinite and most of the results about SUSY breaking etc. would apply. Further all the excited states are always $(p+1)$-fold degenerate. It is amusing to note that in ortho supersymmetricsy QM too [11], the relation between H and charges is exactly as given by eq. (3).

Following the work of [7], we now consider a specific PSQM model of order p which in addition is conformally invariant and show that the conformal PSQM algebra is rather simple. Let us consider the choice

$$W_1=W_2=\cdots=W_p=-\frac{\lambda}{x}. \tag{28}$$

Note that in this case the condition (13) is trivially satisfied when all C_i are zero. The interesting point is that in this case, apart from the p parafermionic charges Q_i, we can also define the dilatation operator D, the conformal operator K and p para superconformal charges S_j so that they form a closed algebra. In particular, on defining

$$D=-\tfrac{1}{4}(xP+Px); \quad K=x^2/2, \tag{29}$$
$$(S_j)_{\alpha\beta}=-x\delta_{\alpha+1,\beta=j+1}, \tag{30}$$

it is easy to show that the algebra satisfied by D, H and K is standard, i.e.

$$[H,K]=2iD, \quad [D,K]=iK, \quad [D,H]=-iH. \tag{31}$$

Further

$$[K,S_j]=0, \quad [H,S_j]=iQ_j, \quad [K,Q_j]=iS_j, \tag{32}$$
$$[D,Q_j]=-\frac{i}{2}Q_j, \quad [D,S_j]=\frac{i}{2}S_j. \tag{33}$$

Besides, apart from the parasusy algebra as described above (with $C_i=0$), we have

$$S_iS_j=0=Q_iS_j=S_iQ_j \quad \text{if} \quad j\neq i+1, \tag{34}$$
$$Q_iS_j^+=0=S_iS_j^+=Q_i^+S_j \quad \text{if} \quad i\neq j, \tag{35}$$

$$S_1 S_1^+ + \sum_{j=1}^{p} S_j^+ S_j = 2K, \tag{36}$$

$$S_1 Q_1^+ + Q_1 S_1^+ + \sum_{j=1}^{p} (S_j^+ Q_j + Q_j^+ S_j) = 4D, \tag{37}$$

It is quite remarkable that an identical algebra also follows in the case of the conformal ortho supersymmetric case [11].

References

[1] L-M Kuang, *J. Phys.* **A26**, L1079 (1991)
[2] D T Pegg and S M Bernett, *Europhys. Lett.* **6**, 483 (1988); *Phys. Rev.* **A39**, 1665 (1989)
X Ma and W Rhodes, *Phys. Rev.* **A43**, 1576 (1991)
[3] J-S Peng and G-X Li, *Phys. Rev.* **A45**, 1289 (1992)
[4] A D Wilson-Gordon, V Buzeck and P L Knight, *Phys. Rev.* **A44**, 7647 (1991)
[5] B Bagchi and P K Roy, *Phys. Lett.* **A200**, 411 (1995)
[6] V A Rubakov and V P Spiridonov, *Mod. Phys. Lett.* **A3**, 1337 (1988)
J Beckers and N Debergh, *Nucl. Phys.* **B340**, 767 (1990)
[7] A Khare, *J. Phys* **A25**, L749 (1992); *J. Math. Phys.* **34**, 1277 (1993)
[8] H A Buchdhal, *Am. J. Phys.* **35**, 210 (1967)
[9] R Kleeman, *J. Aust. Math. Soc.* **B23**, 52 (1981)
[10] H S Green, *Phys. Rev.* **90**, 270 (1953)
Y Ohnuki and S Kamefuchi, *Quantum field theory and parastatistics* (University of Tokyo, 1982)
[11] A Khare, A K Mishra and G Rajasekaran, *Int. J. Mod. Phys.* **7**, 1245 (1993)
[12] For $p = 2, 3$ see B Bagchi and K Samanta, *Multidimensional Parasuperalgebras and a noninteracting N-level system*, unpublished (1992)

physics pp. 205–212

A new class of cosmological models in Lyra geometry*

G P SINGH[1] and KALYANI DESIKAN[2]
Cosmology Group, Department of Mathematics, Indian Institute of Technology, Chennai 600 036, India
Present Address: [1] Department of Mathematics, Vishveshwara Regional Engg. College, Nagpur 440 011, India
[2] Department of Mathematics, M.O.P. Vaishnav College for Women, Nungambakkam, Chennai 600 034, India

MS received 23 April 1996; revised 18 December 1996

Abstract. FRW models have been studied in the cosmological theory based on Lyra's geometry. A new class of exact solutions has been obtained by considering a time dependent displacement field for constant deceleration parameter models of the universe.

Keywords. Cosmology; Lyra geometry; FRW models.

PACS No. 98.80

1. Introduction

Einstein's idea of geometrizing gravitation in the form of general theory of relativity inspired the idea of geometrizing other physical fields. Weyl [1] suggested the first so-called unified theory which is a geometrized theory of gravitation and electromagnetism. But this theory was never taken seriously as it was based on non-integrability of length transfer. Lyra [2] proposed a new modification of Riemannian geometry by introducing a gauge function which removes the non-integrability condition of the length of a vector under parallel transport. In consecutive investigations Sen [3], Sen and Dunn [4] proposed a new scalar-tensor theory of gravitation and constructed an analog of the Einstein field equations based on Lyra's geometry which in normal gauge may be written as

$$R_{ij} - \tfrac{1}{2} g_{ij} R + \tfrac{3}{2} \phi_i \phi_j - \tfrac{3}{4} g_{ij} \phi_k \phi^k = -8\pi G T_{ij}, \tag{1}$$

where ϕ_i is the displacement vector and other symbols have their usual meaning as in Riemannian geometry.

Halford [5] has pointed out that the constant vector displacement field ϕ_i in Lyra's geometry plays the role of cosmological constant Λ in the normal general relativistic treatment. It is shown by Halford [6] that the scalar-tensor treatment based on Lyra's geometry predicts the same effects, within observational limits, as the Einstein's theory. Bhamra [7], Karade and Borikar [8], Kalyanshetti and Wagmode [9], Reddy and Innaiah

*Dedicated to Professor V B Johri on his sixtieth birthday.

[10], Beesham [11], Reddy and Venkateswarlu [12], Soleng [13] have studied cosmology in Lyra's geometry with a constant displacement field. However, this restriction of the displacement field to be constant is merely one of convenience and there is no *a priori* reason for it. Beesham [14] considered FRW models with time dependent displacement field. He has shown that by assuming the energy density of the universe to be equal to its critical value, the models have the $k = -1$ geometry. Singh and Singh [15–18] have presented Bianchi-type I, III and Kantowaski–Sachs cosmological models with time dependent displacement field and have made a comparative study of Robertson–Walker models with constant deceleration parameter in Einstein's theory with cosmological term and in the cosmological theory based on Lyra's geometry. They [18] have given a review of cosmological models in Lyra's geometry.

Most of the well known FRW models of the universe with curvature parameter $k = 0$ are models with constant deceleration parameter. Therefore, in this paper we have studied FRW models with time dependent displacement field and have solved the field equations by taking the deceleration parameter to be constant.

2. Field equations

The time-like displacement vector ϕ_i in (1) is given by

$$\phi_i = (\beta(t), 0, 0, 0). \tag{2}$$

We assume a perfect fluid for which the energy-momentum tensor has the form

$$T_{ij} = (\rho + p)u_i u_j - p g_{ij}. \tag{3}$$

For the FRW metric

$$\mathrm{d}s^2 = \mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\Phi^2)\right],$$

where $k = 1, -1, 0$, the field equations (1) become with the eqs (2) and (3)

$$3H^2 + \frac{3k}{R^2} - \frac{3\beta^2}{4} = \chi\rho, \tag{4}$$

$$2\dot{H} + 3H^2 + \frac{k}{R^2} + \frac{3\beta^2}{4} = -\chi p, \tag{5}$$

where $\chi = 8\pi G$ and $H = \dot{R}/R$ is the Hubble's function. Equations (4) and (5) lead to the continuity equation

$$\chi\dot{\rho} + \tfrac{3}{2}\beta\dot{\beta} + 3[\chi(\rho + p) + \tfrac{3}{2}\beta^2]H = 0. \tag{6}$$

Assuming an equation of state

$$p = \gamma\rho, \qquad 0 \leq \gamma \leq 1 \tag{7}$$

and eliminating $\rho(t)$ from (4) and (5) we have

$$2\dot{H} + 3(1 + \gamma)H^2 + (1 + 3\gamma)\frac{k}{R^2} + (1 - \gamma)\frac{3\beta^2}{4} = 0. \tag{8}$$

Here β^2 plays the role of a variable cosmological term $\Lambda(t)$. We have two independent equations in three unknown viz $R(t)$, $\rho(t)$ and $\beta(t)$. Therefore, we need one more relation among the variables in order to obtain a unique solution. Hence, in what follows we have taken the deceleration parameter to be constant.

3. Solutions of the field equations

Let the deceleration parameter

$$q = -\frac{R\ddot{R}}{\dot{R}^2} = -\frac{(\dot{H}+H^2)}{H^2} = b \text{ (constant)}. \tag{9}$$

The above equation may be rewritten as

$$\frac{\ddot{R}}{R} + b\frac{\dot{R}^2}{R^2} = 0. \tag{10}$$

On integration, equation (10) gives the exact solution

$$R(t) = \begin{cases} (D+Ct)^{1/(1+b)} & b \neq -1, \\ R_0\, e^{H_0 t} & b = -1, \end{cases} \tag{11}$$

where C, D, R_0 and H_0 are constants of integration. Using (9) in (8) leads to

$$\beta^2 = \frac{4}{3(\gamma-1)}\left[(1+3\gamma-2b)H^2 + (1+3\gamma)\frac{k}{R^2}\right]. \tag{12}$$

Now using (12) in (4) gives

$$\chi\rho = \frac{1}{(\gamma-1)}\left[2(b-2)H^2 - \frac{4k}{R^2}\right]. \tag{13}$$

Flat models

With $k = 0$, equations (12) and (13) reduce to

$$\beta^2 = \frac{4}{3(\gamma-1)}(1+3\gamma-2b)H^2 \tag{14}$$

and

$$\chi\rho = \frac{1}{(\gamma-1)}2(b-2)H^2. \tag{15}$$

From (15) we see that $\rho \geq 0$ if $b-2 \leq 0$ i.e. $b \leq 2$ since $(\gamma-1) < 0$. From (14) we see that since $(\gamma-1) < 0$,

$$\beta^2 > 0 \quad \text{if} \quad b > \frac{(1+3\gamma)}{2} \tag{16}$$

and

$$\beta^2 < 0 \quad \text{if} \quad b < \frac{1+3\gamma}{2}. \tag{17}$$

Table 1. Values of β^2 and ρ for dust and radiation power-law models.

γ	β^2	ρ
0	$\frac{4}{3}\frac{(2b-1)}{(1+b)^2t^2}$	$\frac{2(2-b)}{\chi(1+b)^2t^2}$
$\frac{1}{3}$	$\frac{4(b-1)}{(1+b)^2t^2}$	$\frac{3(2-b)}{\chi(1+b)^2t^2}$

Again from equation (14) we see that when $b = (1+3\gamma)/2, \beta^2 = 0$ and the equations reduce to those of the standard FRW flat universe.

Case a: $b \neq -1$. For singular models since $R(0) = 0$, (11) leads to

$$R = R_0 t^{1/(1+b)}. \tag{18}$$

Using (18) in (14) and (15) yields

$$\beta^2 = \frac{4(1+3\gamma-2b)}{3(\gamma-1)}\frac{1}{(1+b)^2t^2} \tag{19}$$

and

$$\chi\rho = \frac{2(b-2)}{(\gamma-1)(1+b)^2}\frac{1}{t^2}. \tag{20}$$

The above expressions for β^2 and energy density $\rho(t)$ are similar to those obtained by Beesham [19] for a variable cosmological term $\Lambda(t)$ and energy density $\rho(t)$. Here β^2 plays the role of a variable cosmological term Λ.

Equations (7) and (20) yield

$$\rho + 3p = \frac{2(1+3\gamma)(b-2)}{(\gamma-1)(1+b)^2}\frac{1}{t^2}.$$

It can be seen from the above expression that the condition $\rho + 3p \geq 0$ would hold only for $1 + 3\gamma \geq 0$ i.e. $\gamma \geq -1/3$. So for values of $\gamma < -1/3$ we cannot have viable models.

We observe from (19) and (20) that β^2 and ρ fall off as $1/t^2$ irrespective of the equation of state. The expressions for β^2 and ρ corresponding to $\gamma = 0, 1/3$ are summarized in the table 1.

It can be readily seen from (19) and (20) that the expressions for β^2 and ρ will not be valid for the empty universe (i.e. $p = \rho = 0$) and the stiff matter (i.e. $p = \rho$) models. We shall now discuss these models.

Empty universe

In the case of empty universe ($p = \rho = 0$) eqs (4) and (5), with $k = 0$, now become

$$3H^2 - \frac{3\beta^2}{4} = 0, \tag{21}$$

$$2\dot{H} + 3H^2 + \frac{3\beta^2}{4} = 0. \tag{22}$$

Adding equations (21) and (22) yields

$$\dot{H} + 3H^2 = 0. \tag{23}$$

On comparing (9) and (23) we get $b = 2$. Therefore, (18) reduces to

$$R = R_0 t^{1/3}. \tag{24}$$

Using (24) in (21) leads to

$$\beta^2 = \frac{4}{9t^2}. \tag{25}$$

From (25) it can be seen that $\beta^2 > 0$ for all values of the deceleration parameter.

Stiff matter model

In the case of Zeldovich fluid viz., $p = \rho$ (i.e. $\gamma = 1$) equations (4), (5) and (8), with $k = 0$, reduce to

$$3H^2 - \frac{3\beta^2}{4} = \chi\rho, \tag{26}$$

$$2\dot{H} + 3H^2 + \frac{3\beta^2}{4} = -\chi\rho, \tag{27}$$

$$\dot{H} + 3H^2 = 0. \tag{28}$$

From (26) it can be seen that $\rho \geq 0$ and $\rho + 3p \geq 0$ for $\beta^2 \leq 0$ since $H^2 \geq 0$ and for $0 \leq \beta^2 \leq 4H^2$.

On comparing equations (9) and (28) we see that $b = 2$. Therefore, (18) reduces to

$$R = R_0 t^{1/3}. \tag{29}$$

Here the expressions for β^2 and ρ cannot be determined uniquely.

Case b: $b = -1$. In this case, equation (9) becomes

$$\dot{H} = 0 \quad \text{and} \quad H = H_0 = \text{constant}. \tag{30}$$

Using (30) in (14) and (15) we have

$$\beta^2 = \frac{4(1+\gamma)}{(\gamma-1)} H_0^2 = \text{constant} \tag{31}$$

and

$$\chi\rho = \frac{-6H_0^2}{(\gamma-1)} = \text{constant}. \tag{32}$$

From (31) it can be seen that since $(1+\gamma) > 0$ and $(\gamma - 1) < 0, \beta^2 < 0$ and from (32) we have $\rho > 0$ for all times since $(\gamma - 1) < 0$.

The expression for β^2 given by (31) is similar to the expression for the cosmological term Λ given by (32) with $k = 0$ in ref. [19]. From (7) and (32) we have

$$\rho + 3p = \frac{6(1+3\gamma)H_0^2}{(1-\gamma)}.$$

Table 2. Values of β^2 and ρ for dust and radiation exponenti models.

γ	β^2	ρ
0	$-4H_0^2$	$\frac{6H_0^2}{\chi}$
$\frac{1}{3}$	$-8H_0^2$	$\frac{9H_0^2}{\chi}$

As in the previous case, we have $\rho + 3p \geq 0$ only for $1 + 3\gamma \geq$ expressions for β^2 and ρ corresponding to $\gamma = 0, 1/3$ are given in t

For the empty universe and stiff matter models the expressions fo determined.

In the next section we shall discuss models with non-zero curvat

Non-flat models

Case a: $b \neq -1$. Using (18) in (12) and (13) yields

$$\beta^2 = \frac{4}{3(\gamma-1)}\left[\frac{(1+3\gamma-2b)}{(1+b)^2t^2} + \frac{(1+3\gamma)k}{R_0^2 t^{2/(1+b)}}\right]$$

and

$$\chi\rho = \frac{1}{(\gamma-1)}\left[\frac{2(b-2)}{(1+b)^2}\frac{1}{t^2} - \frac{4k}{R_0^2 t^{2/(1+b)}}\right].$$

An obvious requirement that must be imposed is $\rho \geq 0$, which can $k \geq 0$ with $b - 2 \leq 0$ (i.e. $b \leq 2$) as $(\gamma - 1) < 0$. The case $k =$ discussed in the previous section.

The above expressions for ρ and β^2 are similar to those of ρ and (45) and (46) of ref. [19].

and

$$\chi\rho = \frac{1}{(\gamma-1)}\left[\frac{2(b-2)}{(1+b)^2t^2} - \frac{4}{R_0^2 t^{2/(1+b)}}\right].$$

From eq. (35), we see that for $b < (1+3\gamma)/2$, $\beta^2 < 0$ fo $(\gamma-1) < 0$.

Again from equation (35) we observe that for $(1+3\gamma)/2 < b$

$$0 < t^{2b/(1+b)} < \frac{(2b-1-3\gamma)R_0^2}{(1+3\gamma)(1+b)^2}$$

and $\beta^2 < 0$ for

$$t^{2b/(1+b)} > \frac{(2b-1-3\gamma)R_0^2}{(1+3\gamma)(1+b)^2}.$$

At

$$t^{2b/(1+b)} = \frac{(2b-1-3\gamma)R_0^2}{(1+3\gamma)(1+b)^2}, \quad \beta^2 = 0.$$

When $b = (1+3\gamma)/2$, equations (35) and (36) reduce to

$$\beta^2 = \frac{4(1+3\gamma)}{3(\gamma-1)R_0^2 t^{4/3(1+\gamma)}}$$

and

$$\chi\rho = \frac{1}{(1-\gamma)}\left[\frac{3(1-\gamma)}{(1+b)^2t^2} - \frac{4}{R_0^2 t^{4/3(1+\gamma)}}\right].$$

From equation (37), it is obvious that $\beta^2 < 0$ for all times expressions for β^2 and ρ cannot be determined for the empty un and stiff matter $(p = \rho)$ models.

Case b: $b = -1$. Using (30) in (12) and (13) leads to

The expression for β^2 given by equation (39) is similar to th cosmological term given by equation (52) of ref. [19].

As in the previous case we have $\rho + 3p \geq 0$ only for $\gamma \geq -\frac{1}{3}$. Fo we cannot have viable models.

For the empty universe and stiff matter models the expressions f determined.

4. Conclusion

In this paper we have obtained exact solutions of Sen equations constant deceleration parameter. The behaviour of the displacer energy density ρ have been examined for both the (i) power-law expansions of both the flat universe and the non-flat universe.

Acknowledgements

One of the authors GPS acknowledges with thanks financial s Department of Atomic Energy, India. KD would like to thank DS support and S Gunasekaran for his invaluable help in typing the n

References

[1] H Weyl, *Sber Preuss. Akad. Wiss.* 465 (1918)
[2] G Lyra, *Math. Z.* **54**, 52 (1951)
[3] D K Sen, *Z. Phys.* **149**, 311 (1957)
[4] D K Sen and K A Dunn, *J. Math. Phys.* **12**, 578 (1971)
[5] W D Halford, *Austr. J. Phys.* **23**, 863 (1970)
[6] W D Halford, *J. Math. Phys.* **13**, 1399 (1972)
[7] K S Bhamra, *Austr. J. Phys.* **27**, 541 (1974)
[8] T M Karade and S M Borikar, *Gen. Relativ. Gravit.*, **9**, 431 (1978)
[9] S B Kalyanshetti and B B Waghmode, *Gen. Relativ. Gravit.* **14**, 82
[10] D R K Reddy and P Innaiah, *Astrophys. Space Sci.* **123**, 49 (1986)
[11] A Beesham, *Astrophys. Space Sci.* **127**, 189 (1986)

PRAMANA
— journal of
physics

© Printed in India

Inhomogeneous cosmological models with heat f

L K PATEL[1], RAMESH TIKEKAR[2] and NARESH DADHICH
Inter-University Centre for Astronomy and Astrophysics, Post Bag 4,
411 007, India
[1]Permanent Address: Department of Mathematics, Gujarat University,
India
[2]Permanent Address: Department of Mathematics, Sardar Patel
Vidyanagar 388 120, India
e-mail: root@patel.ernet.in
naresh@iucaa.ernet.in

MS received 6 November 1995; revised 27 September 1996

Abstract. We present a general class of inhomogeneous cosmologic
non-thermalized perfect fluid by assuming that the background spacetime
commuting Killing vectors and has separable metric coefficients. The si
these models depends on the choice of the parameters and the metric fu
previously known perfect fluid models follow as particular cases of this g
and geometrical features of these models are studied and the general expr
distribution is given.

Keywords. Inhomogeneous cosmology models; heat flux.

PACS Nos 04.20; 98.80

1. Introduction

The standard Friedman–Robertson Walker (FRW) cosmologica
scribes a homogeneous and isotropic distribution for its matter co
successful in describing the present state of the universe. It is thou
homogeneous and isotropic character of the spacetime cannot be s

Dadhich, Tikekar and Patel [3] have established a link between tl
the singularity free family by deducing the latter through a
inhomogenization and anisotropization of the former. The matter
models has been a perfect fluid in thermal equilibrium. At very earl
universe is supposed to be in a highly dense and hot state. Subseq
of dissipative effects in cosmology becomes highly significant. I
cosmology were considered in the context of large entropy per ba
microwave background radiation by Misner [4] and Caderni a
mological models with heat flux have been studied by several aut

In this paper we have considered an inhomogeneous metric
space-like commuting Killing vectors and has separable metric
tained a general class of inhomogeneous solutions of Einstein's
non-thermalized perfect fluid as the source term. Explicit exact solu
particular cases have been obtained. These particular solutions i
generalizations of the general class of inhomogeneous perfect flu
and Senovilla [2].

2. Metric and field equations

We begin with the spacetime which admits two commuting space
fields that are hyper-surface orthogonal. It is well known that
spacetimes can be put in the generalized Einstein–Rosen form. I
that the metric coefficients are separable functions. Ruiz and Se
metric in the form

$$ds^2 = T^{2m}F^2(dt^2 - H^2dx^2) - T^{n+1}GPdy^2 - T^{1-n}GP^{-}$$

where $T = T(t)$ and F, G, P, H are functions of only x, m and n
freedom of choosing co-ordinates is used in having the same tim
metric coefficients g_{11} and g_{44}. The co-ordinates are labelled as x
and $x^4 = t$. The function $H(x)$ is introduced to facilitate the integra
equations in terms of elementary functions. Note that the equatio
not been chosen *a priori*. Further one does not even have to assume
come from the single function $T(t)$, it has been shown [17] that

$$R_{22}T^{2m}F^2 = \frac{1}{H^2}\left[\frac{1}{2}\left(\frac{G''}{G} - \frac{P''}{P}\right) + \frac{1}{2}\frac{P'^2}{P^2} - \frac{1}{2}\frac{G'P'}{GP} - \frac{1}{2}\right.$$

$$-\frac{1}{2}(1+n)\frac{\ddot{T}}{T},$$

$$R_{33}T^{2m}F^2 = \frac{1}{H^2}\left[\frac{1}{2}\left(\frac{G''}{G} + \frac{P''}{P}\right) - \frac{1}{2}\frac{P'^2}{P^2} + \frac{1}{2}\frac{G'P'}{GP} - \frac{1}{2}\right.$$

$$-\frac{1}{2}(1-n)\frac{\ddot{T}}{T},$$

$$R_{44}T^{2m}F^2 = -\frac{1}{H^2}\left[\frac{F''}{F} - \frac{F'^2}{F^2} - \frac{F'H'}{FH} + \frac{F'G'}{FG}\right] + (m -$$

$$+\frac{1}{2}(n^2 - 1 - 4m)\frac{\dot{T}^2}{T^2},$$

$$R_{14}T^{2m}F^2 = \frac{1}{2H}\frac{\dot{T}}{T}\left[(1-2m)\frac{G'}{G} + n\frac{P'}{P} - 2\frac{F'}{F}\right].$$

Here and in what follows, a prime and a dot indicate derivativ
t respectively.

The matter content of the spacetime is assumed to be a non-t
described by the energy momentum tensor

$$T_{\alpha\beta} = (\rho + p)U_\alpha U_\beta - pg_{\alpha\beta} + (q_\alpha U_\beta + q_\beta U_\alpha),$$

$$U_\alpha U^\alpha = 1 \quad U_\alpha q^\alpha = 0,$$

where ρ and p denote the energy density and fluid pressure of
denotes unit time-like flow vector of fluid and q^α denotes the spa
orthogonal to U^α. The components of these vectors in the tetrad f

$$U^\alpha = (0,0,0,1), \quad q_\alpha = (q,0,0,0)$$

where q is a function of co-ordinates governed by the Einstein

where d and λ are constants. The equations (11) written out explic

$$n\frac{\ddot{T}}{T}=\frac{1}{H^2}\left[\frac{P''}{P}-\frac{P'^2}{P^2}+\frac{G'P'}{GP}-\frac{P'H'}{PH}\right]$$

$$\frac{\ddot{T}}{T}(1-n-2m)=\frac{1}{H^2}\left[-\frac{G''}{G}-\frac{P''}{P}+\frac{G'^2}{G^2}-\frac{2F''}{F}-\frac{G'P'}{GP}+\right.$$

$$\left.+\frac{H'G'}{HG}+\frac{P'H'}{PH}+\frac{2F'^2}{F^2}+2\frac{F'H'}{FH}\right].$$

The left hand sides of these equations are functions of time only v sides are functions of x only. Therefore both the sides of respectiv equal to separation constants (different for different equations). T

$$\frac{\ddot{T}}{T}=\varepsilon a^2,\ \varepsilon=0,\ \pm 1,\ a=\text{constant}$$

implying

$$T(t)=A\cosh(at)+B\sinh(at),\ \varepsilon=1$$

$$T(t)=At+B,\ \varepsilon=0$$

$$T(t)=A\cos(at)+B\sin(at),\ \varepsilon=-1,$$

where A and B are constants of integration. On using (15) and (18 (17) take the forms

$$\alpha'+\alpha\beta-\alpha\frac{H'}{H}=\varepsilon n a^2 H^2$$

and

$$(2d+1)\beta'-2d\beta^2-(2d+1)\beta\frac{H'}{H}+\alpha^2-4\lambda\alpha\beta=(2m-$$

where

$$\alpha=\frac{P'}{P},\ \beta=\frac{G'}{G}$$

where

$$Z = (1+4d)\beta^2 - \alpha^2 + 4\lambda\alpha\beta.$$

In the above, it is assumed that $2d+1 \neq 0$. From equations (23) and (24) it e
that

$$\rho = \frac{2d-3}{2d+1}p + \frac{1}{16\pi T^{2m}F^2}\left[\frac{2(4m+1-n^2)}{2d+1}\left(\frac{\dot{T}^2}{T^2} - \varepsilon a^2\right)\right.$$
$$\left. + \frac{\varepsilon a^2(2d+2m-1+2\lambda n-n^2)}{2(2d+1)}\right].$$

Accordingly the matter distribution will satisfy an equation of state $\rho =$
one of the following conditions:

(i) $\dot{T}^2/T^2 = \varepsilon a^2$, $2d+2m-1+2\lambda n-n^2 = 0$,
(ii) $4m+1 = n^2$, $2d+2m-1+2\lambda n-n^2 = 0$,
(iii) $\varepsilon = 0$, $n^2 = 4m+1$,

where $k = (2d-3)/(2d+1) \neq 1$. However, the parameters can be chosen app
that the matter distribution satisfies equation of state $\rho = p$ corresponding t

The parameter q which implies the presence of heat flow with matter
any one of the following conditions is fulfilled:

(i) $2d = 1-2m$, $n = 2\lambda$
(ii) $\alpha = 0$, $\beta = 0$
(iii) $\beta = 0$, $n = 2\lambda$
(iv) $\alpha = 0$, $2d = 1-2m$.

Accordingly the set up developed above is used to obtain heat flow general
general class of inhomogeneous perfect fluid solutions of Ruiz and Se
$\alpha \neq 0$, $\beta \neq 0$, and $2d \neq 1-2m$ or $n \neq 2\lambda$.

We give below the expressions for the expansion scalar θ, the shear sca
acceleration vector f_i describing the kinematical behaviour of the velocity
by (9)

$$\theta = \frac{(m+1)\dot{T}}{}$$

where ψ is the thermal conductivity and $\mathscr{T}$ is the temperature. If th
ψ is a function of time alone, then the above equation can be formal

$$\mathscr{T} = -\frac{F\dot{T}}{16\pi\psi T^{m+1}}\int F^{-2}\{(1-2m-2d)\alpha+(n-2\lambda)\beta\}$$

where $l_0(t)$ is an arbitrary function of time.

In the next section explicit solutions of the differential equatio
certain special cases are obtained.

3. Explicit solutions

We have not made any attempt to obtain the general solutio
equations (20) and (21). It is observed that these equations admit s
elementary functions in certain special cases. We discuss brie
omitting the details of calculations.

Case 1. The simplest family of solutions is characterized by $\varepsilon = 0$

$$T(t) = At + B.$$

In this case (20)–(22) can be integrated completely if $4\lambda^2 + 4d + 1$
solution is obtained in the form

$$P = (N + C^\gamma)^{l_1}, G = P^{l_2}C^{l_3}, H = MC'P^{l_4},$$

where C is an arbitrary function of x, M and N are arbitrary cons

$$l_2 = (4d+1)^{-1}(-2\lambda \pm \sqrt{4\lambda^2+4d+1}), l_3 = -(2d+1)$$

$$l_4 = (1+2d)^{-1}[-4\lambda-(1+6d)l_2], l_1 = (1+2d)[2(1+$$

$$\gamma = (1+4d)/4d.$$

The physical variables are given by

$$16\pi\rho T^{2m}F^2 = \frac{A^2(4m+1-n^2)}{2(At+B)^2} - \frac{(2d+1)(4d+1)(2d-}{16d^2M^2P^{2l_4}C^2}$$

$d = -\lambda \neq -3/2$, the above solution represents heat flux counterpart of universe discussed by Davidson[18].

The solutions of equations (20)–(22) for the exceptional cases $d =$ $d = -\frac{1}{4}$ and $4\lambda^2 + 4d + 1 = 0$ are given below:

Case $d = 0$: $P = (N + C)^{l_1}$, $G = P^{l_2} C^{-1}$, $H = MC' P^{l_4} C^{-1}$ where M, N constants and

$$l_1 = l_2/(1 + l_2^2),\ l_4 = -4\lambda - l_2,\ l_2^2 = \tfrac{1}{2}[-1 - 4\lambda + \sqrt{(1 - 4\lambda)^2}$$

with $\lambda < 1/4$.

Case $d = -\frac{1}{2}$: $P = A_1 G^l$, $G = (2\lambda \pm \sqrt{4\lambda^2 - 1}) K_1 C + K_2$, $H = C'$,

$l(2\lambda \pm \sqrt{4\lambda^2 - 1}) = 1$, where A_1, K_1, K_2 are constants and λ^2

Case $d = -\frac{1}{4}$: $P = (ax + b)^{1/8\lambda}$, $G = \mathrm{e}^{-ax} P^{1/4\lambda}$, $H = \mathrm{e}^{-ax} P^{(1 - 32\lambda^2)/4\lambda}$. Here arbitrary constants.

Case $4\lambda^2 + 4d + 1 = 0$: $P = \exp(ax^{8\lambda^2/(1 + 4\lambda^2)})$ $G = P^{1/2\lambda}(ax)^{(1 - 4\lambda^2)/(1 + 4\lambda^2)}$, $H =$

In the absence of heat flux the solutions corresponding to the cases $d =$ and $4\lambda^2 + 4d + 1 = 0$ respectively reduce to the fluid solutions corresp cases $m = 3/4$, $m = 1/2$ and $4m = n^2 + 3$ discussed by Ruiz and Senov $d = -1/2$ is a new case.

Case 2. For $\varepsilon \neq 0$, there are several integrable cases. We discuss a solution $\varepsilon = 1$ in brief. The solution has

$$H = 1,\ P = \mathrm{e}^{-l_1 \mu a x} f^{l_1},\ G = e^{\mu a x} f,$$

$$f = A_1 e^{\max} + A_2 e^{-\max},\ l_1 = \sqrt{-2d},\ m^2 = \frac{n + l_1 \mu^2}{l_1}$$

$$\mu^2 = \frac{1}{4(1 + 2d)}\left[2(2m - 1) - \frac{n}{\sqrt{-2d}}(3 + 10d)\right],\ \lambda^2 = -(1 + 6$$

$$32\pi pT^{2m}F^2 = a^2(4m-n^2-3)+(4m+1-n^2)\left(\frac{\dot{T}^2}{T^2}-\right.$$

$$+\left[a^2\mu^2\{1+6d-\sqrt{1+6d}\}+\frac{f'^2}{f^2}\{1+\right.$$

$$\left.+2a\mu\frac{f'}{f}(1+2d)\right],$$

$$16\pi qT^{2m}F^2 = -\frac{\dot{T}}{T}\left[\frac{f'}{f}\left\{1-2m-2d+\sqrt{-2d}\left(n-\frac{\sqrt{}}{2\sqrt{}}\right.\right.\right.$$

$$\left.+\mu a\left\{1-2m-2d-\sqrt{-2d}\left(n-\frac{\sqrt{1+}}{2\sqrt{-}}\right.\right.\right.$$

This solution is the heat flow generalization of the perfect flu
obtained by Ruiz and Senovilla.

Case 3. This case is characterized by

$$\varepsilon = \pm 1,\ 1+4d+4\lambda^2>0,\ d\neq -\tfrac{1}{2},\ d\neq +\tfrac{1}{4}$$

$$\frac{n(1+2d)(1-4d)}{2m+2n\lambda-1+8n\lambda d} = 2\lambda \pm \sqrt{1+4d+4\lambda^2}.$$

An explicit solution obtained in this case is

$$H=1,\ P=f^{l_1},\ G=ff''^{l_2}$$

where $l_1=2\lambda\pm\sqrt{1+4d+4\lambda^2}$, $l_2=(1+2d)/2d$ and $f''/f=-(2\varepsilon na^2d/($
The physical parameters ρ, p and q are given by

$$32\pi\rho T^{2m}F^2 = \varepsilon a^2\left[2m+2-n^2+2\lambda n-\frac{2d-3}{2d+1}\{1-2n\right.$$

$$\left.+4ndl_2(l_1-4\lambda)\}\right]+(4m+1-n^2)\left(\frac{\dot{T}^2}{T^2}\right.$$

This solution is the heat flow generalization of the perfect fluid case-3-solu
Ruiz and Senovilla. The function f will be in the form of trigonometric
function depending on the value of ε.

Case 4. Here we will obtain the heat flow generalization of the case-4 flu
Ruiz and Senovilla characterized by the relation $2m = 1 + n$. The metri
this solution are given by

$$H = \frac{MC'}{P},\ G = PC^{(-1+2d)/2d},\ P^2 = -2\varepsilon nda^2M^2C^2 + NC^{(4d+1)/2}$$

together with

$$2m = n + 1,\ d + \lambda = 0.$$

The density, pressure and heat flow parameter are given as

$$32\pi\rho T^{2m}F^2 = \varepsilon a^2\left[4(1-m^2) - \frac{(2m-1)}{2d}(2d-3)(2d+1)\right]$$

$$+ 4m(2-m)\left(\frac{\dot{T}^2}{T^2} - \varepsilon a^2\right) - \frac{(2d+1)(4d+1)(2d}{4d^2}$$

$$32\pi pT^{2m}F^2 = -\varepsilon a^2\left[4(1-m^2) + \frac{n}{2d}(2d+1)^2\right] + 4m(2-m)\Big($$

$$-\frac{(2d+1)^2(4d+1)Q}{4d^2},$$

$$32\pi qT^{2m}F^2 = (2d+1)(1-2m-2d)P\dot{T}/MT\mathrm{d}C.$$

The metric of the spacetime of this class of solutions has well defin
symmetry and it contains a large subclass of singularity free cosmologica
heat flow [17].

$$32\pi pT^{2m}F^2 = 4a^2(1-m^2) + 4m(2-m)\left(\frac{\dot{T}^2}{T^2} + a^2\right)$$

$$+ 4(1+2d)^2\frac{na^2}{8d},$$

$$8\pi qT^{2m}F^2 = -(1+2d)(1-2m-2d)\frac{\dot{T}}{T}\frac{f'}{f}.$$

The other type of solutions has

$$H = 1,\ G = f^{-(1+2d)/2d}f',\ P = f',\ \frac{f''}{f} = 2na^2d < 0.$$

The physical variables have the explicit expressions

$$32\pi\rho T^{2m}F^2 = -4a^2(1-m^2) + 4m(2-m)\left(\frac{\dot{T}^2}{T^2} + a^2\right)$$

$$+ (2d-3)(2d+1)\left[\frac{1+4d}{4d^2}\frac{f'^2}{f^2} - 2(2m-$$

$$32\pi pT^{2m}F^2 = 4a^2(1-m^2) + 4m(2-m)\left(\frac{\dot{T}^2}{T^2} + a^2\right)$$

$$+ (1+2d)^2\left[\frac{1+4d}{4d^2}\frac{f'^2}{f^2} - 2(2m-1)a^2\right.$$

$$8\pi qT^{2m}F^2 = \frac{(1+2d)}{4d}(1-2m-2d)\frac{\dot{T}}{T}\frac{f'}{f}.$$

The former type of solution includes a heat flow generalization symmetric stiff fluid solution of Davidson [19]. When the heat fl stipulating $2d = 1 - 2m$, the solutions reduce to inhomogeneous c ric perfect fluid solutions other than those obtained by Ruiz and

In the next section we state in brief the physical and geomet solutions reported above and deduce the singularity-free subfam

However a number of solutions admit a linear equation of state $p =$ Ruiz–Senovilla class of solutions. In each case of solutions the expr ture distribution can be obtained by using the respective explicit exp variables P, G and T in equation (30).

The subfamily of solutions (46), (47) of case-4 defined by

$$\varepsilon = 1,\ n \geqslant 3,\ T(t) = \cosh(at),\ C(x) = \cosh(kax),$$

where k is a real constant, provides singularity free models with element of the spacetime for these models can be expressed in the

$$ds^2 = \cosh^{2m}(at)[\cosh(kar)]^{-(1+2d)}\left[dt^2 - \frac{M^2k^2a^2}{P^2}\sinh^2\right.$$
$$- [\cosh(at)]^{2-2m}[\cosh(kar)]^{-(1+2d)/2d}dz^2$$
$$- [\cosh(at)]^{2m}[\cosh(kar)]^{-(1+2d)/2d}\cdot P^2 d\phi^2,$$

where

$$P^2 = 2(1-2m)da^2M^2\cosh^2(kar) + N[\cosh(kar)]^{(4d+1)/2}$$

Here the co-ordinates x and y are replaced by r and ϕ. The rang taken to be $-\infty < T, z < \infty, 0 \leqslant r < \infty, 0 \leqslant \phi \leqslant 2\pi$. The density, flow parameter for this class of solutions have the following expre

$$32\pi\rho[\cosh(at)]^{2m}[\cosh(kar)]^{-(1+2d)}$$
$$= a^2\left[4(1-m^2) - \frac{(2m-1}{2d}\right.$$
$$+ 4m(m-2)a^2\operatorname{sech}^2(at$$
$$- \frac{(2d+1)(4d+1)(2d-}{4d^2}$$

$$32\pi p[\cosh(at)]^{2m}[\cosh(kar)]^{-(1+2d)} = -a^2\left[4(1-m^2) -\right.$$
$$+ 4m(m-2)\operatorname{sech}$$

References

[1] J M M Senovilla, *Phys. Rev. Lett.* **64**, 2219 (1990)
[2] E Ruiz and J M M Senovilla, *Phys. Rev.* **D45**, 1995 (1992)
[3] N Dadhich, R Tikekar and L K Patel, *Curr. Sci.* **65**, 694 (1993)
[4] C W Misner, *Astrophys. J.* **151**, 431 (1968)
[5] N Caderni and R Fabri, *Nuovo Cimento* **B44**, 288 (1978)
[6] Y Dang, *Gen. Relativ. Gravit.* **21**, 503 (1989)
[7] G Mukherjee, *J. Astrophys. Astron.* 7, 259 (1986)
[8] M Novello and M J Reboucas, *Astrophys. J.* **225**, 719 (1978)
[9] D Ray, *J. Math. Phys* **21**, 2797 (1980)
[10] M J Reboucas and J A S de Limma, *J. Math. Phys.* **23**, 1547 (1982)
[11] M J Reboucas and J A S de Limma, *J. Math. Phys.* **22**, 2699 (1981)
[12] M J Reboucas, *Nuovo Cimento* **B67**, 120 (1980)
[13] J M Bradley and E Sveistins, *Gen. Relativ. Gravit.* **16**, 1119 (1984)
[14] E Sveistins, *Gen. Relativ. Gravit.* **17**, 521 (1985)
[15] L K Patel and N Dadhich, *Astrophys. J.* **401**, 433 (1991)
[16] L K Patel and N Dadhich, *Class. Quantum Gravit.* **10**, L85 (1993)
[17] N Dadhich, L K Patel and R Tikekar, *Pramana–J. Phys.* **44**, 303 (1995
[18] W Davidson, *J. Math. Phys.* **32**, 1560 (1991)
[19] W Davidson, *Class. Quantum Gravit.* **10**, 1843 (1993)

PRAMANA
—journal of
physics

© Printed in India

Photon production from baryon rich quark-glu

PRAGYA, ASHOK GOYAL and J D ANAND
Department of Physics and Astrophysics, University of Delhi, Delhi 110 007,

MS received 16 October 1996; revised 21 March 1997

Abstract. We study high energy photon production from a quark-gluon plas density. We find that the photon production spectrum from the quark-gluon pl constant temperature is only mildly dependent on the quark chemical potentia

Keywords. Phase transition; quark-gluon plasma; photon production; heavy

PACS Nos 25.75; 12.38; 14.80

1. Introduction

The ongoing and planned experiments involving ultrarelativistic heavy believed to unravel the exciting possibility of a transition from normal n deconfined phase of quarks and gluons (QGP). Among the various formation are the electromagnetic 'probes', both virtual and real pl particularly suitable to carry information of the QGP state since they ar reaction volume with little interaction and hence act as ideal markers o history of the QGP state. The production rates of dileptons (from virtu real high-energy photons from QGP have been calculated in the literatu been compared with their production rates from the hot hadronic matte at the definitive conclusion regarding the formation of QGP. A detaile photon production at $T = 200$ MeV led Kapusta, Lichard and Seibert [3 'the hadron gas shines just as brightly as the quark-gluon plasm

also confirmed in numerical simulations of nuclear collisions e quantum molecular dynamics [6] and a relativistic cascade model [7 at RHIC and at CERN's anticipated LHC, recent calculations [8 energies $\sim$ 100–200 GeV per nucleon, the colliding heavy ion transparent and the baryonic chemical potential may not be sm temperature, at least in the experimentally accessible region of ra proposed experiments are unable to measure the electromagnetic baryon free mid rapidity ($y = 0$) region.

Thus with the possibility that baryon density may be larg thermodynamic equilibrium is a function of both the temperatu potential μ_B. This would lead to modification of the dilepton and p Recent calculations by Dumitru *et al* [9] indeed demonstrated such that the dilepton producton rate is a strongly decreasing function μ/T estimated from heavy ion collisions lying between 1 and 2, th production suppressed by a factor 3–10, depending on the value mass and momentum, the plasma energy density etc. In this co to estimate the effect of finite chemical potential on the producti coming from the $q\bar{q}$ annihilation process $q\bar{q} \rightarrow g\gamma$ and from the QC $qg \rightarrow q\gamma, \bar{q}g \rightarrow \bar{q}\gamma$. Direct photon production in the hadronic sector the pseudo scalar and vector meson decays however, will not be affe of the chemical potential.

2. Photon production rate

Consider the annihilation process

$$q(p_1) + \bar{q}(p_2) \rightarrow g(p_3) + \gamma(p)$$

and the Compton processes

$$q(p_1)(\bar{q}(p_1)) + g(p_2) \rightarrow q(p_3)(\bar{q}(p_3)) + \gamma(p)$$

and let $\sum |M_i|^2$ be the matrix element squared and summed over f degrees of freedom. The photon production rate per unit volume is

Defining $s = (p_1 + p_2)^2$ and $t = (p_1 - p)^2$ and converting the total ra rate using the δ function, all the integrals can be done except those ov

$$E\frac{\mathrm{d}R_i}{\mathrm{d}^3p} = \frac{N_i}{2(2\pi)^6}\frac{T}{32E}\left[\frac{1}{\mathrm{e}^{(E-\mu)/T}}\frac{1}{1+\mathrm{e}^{\mu/T}}\right.$$
$$\times \int \frac{\mathrm{d}s}{s}\{\ln(1+\mathrm{e}^{-((s/4E)+\mu)/T}) - \ln(1-\mathrm{e}^{-s/4ET}$$
$$\left.\times \int \mathrm{d}t \sum |M_i|^2 + \mu \leftrightarrow -\mu\right].$$

For relativistic particles the matrix element square is related to th section by [3]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{\Sigma|M|^2}{16\pi s^2}.$$

For the annihilation process, in this limit

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{8\pi\alpha\alpha_c}{9s^2}\frac{u^2+t^2}{ut}$$

and $N = 20$ when we sum over u and d quarks.

For the two Compton processes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{-\pi\alpha\alpha_c}{3s^2}\frac{u^2+s^2}{us}$$

and $N = 320/3$.

The integral over t gives the total cross-section and if the quark n the cross-section would diverge. A non-vanishing quark mass regu Although the bare quark mass is very small, the effective mass interacting thermal quark-gluon plasma is considerably modified d effects of the medium at high temperatures and densities. Braaten an that a quark acquires an effective thermal mass $m_{\mathrm{th}}^2 = 2\pi\alpha_c/3(T^2 +$ *et al* [3] showed that the net effect of the medium on photon produc

where

$$I_a = \int_{k_c^2/ET}^{\infty} \ln(1+\mathrm{e}^{-x+\mu/T})\left\{\ln\left(\frac{4ETx}{k_c^2}\right)-1\right\}\mathrm{d}x \text{ and } C$$

$$E\frac{\mathrm{d}R_{\mathrm{comp.}}}{\mathrm{d}^3p} = \frac{5}{9}\frac{T^2}{\pi^4}\alpha\alpha_c\left[\frac{\pi^2}{6}\left\{\ln\left(\frac{4ET}{k_c^2}\right)+C_B'\right\}\right.$$
$$\times\left\{\frac{1}{\mathrm{e}^{-\mu/T}+1}\frac{1}{\mathrm{e}^{(E+\mu)/T}+1}+(\mu\right.$$
$$+\frac{1}{1+\mathrm{e}^{-\mu/T}}\frac{1}{\mathrm{e}^{(E+\mu)/T}+1}\sum_{n=1}^{\infty}(-$$
$$\times\left(\ln\left(\frac{4ET}{k_c^2}\right)-C_{\mathrm{Euler}}+\frac{1}{2}-\ln\right.$$
$$\left.\left.+\frac{1}{1+\mathrm{e}^{\mu/T}}\frac{1}{\mathrm{e}^{(E-\mu)/T}+1}I_c\right],\right.$$

where

$$I_c = \int_{k_c^2/4ET}^{\infty} \ln(1+\mathrm{e}^{-x+\mu/T})\left\{\ln\left(\frac{4ETx}{k_c^2}\right)+\frac{1}{2}\right\}\mathrm{d}x,$$

$$C_B' = -C_{\mathrm{Euler}}+\frac{1}{2}-\frac{6}{\pi^2}\sum\frac{\ln n}{n^2} = -0.6472.$$

For baryon free plasma $\mu = 0$ and the integrals I_a and I_c can be p
and we recover the well-known results [3] namely,

$$E\frac{\mathrm{d}R_{\mathrm{comp.}}}{\mathrm{d}^3p} = \frac{5}{9}\frac{\alpha\alpha_c}{6\pi^2}T^2\mathrm{e}^{-E/T}\left\{\ln\left(\frac{4ET}{k_c^2}\right)+C_F\right\},$$

$$E\frac{\mathrm{d}R_{\mathrm{ann.}}}{\mathrm{d}^3p} = \frac{5}{9}\frac{\alpha\alpha_c}{3\pi^2}T^2\mathrm{e}^{-E/T}\left\{\ln\left(\frac{4ET}{k_c^2}\right)+C_B\right\}.$$

Table 1. Photon thermal production rate $E(dR/d^3p)$ in fm^{-4} μ/T for $E = 1, 2$ and 3 GeV at two different quark-gluo $T = 200$ MeV and 150 MeV respectively.

	$T = 200$ MeV			$T =$
Photon energy E	1 GeV	2 GeV	3 GeV	1 GeV
μ/T	Rate $\times 10^5$	Rate $\times 10^7$	Rate $\times 10^9$	Rate $\times 10^6$
0.0	1.734	1.817	1.478	2.270
0.5	1.631	1.732	1.414	2.151
1.0	1.585	1.707	1.398	2.109
1.5	1.575	1.725	1.416	2.121
2.0	1.591	1.785	1.468	2.181
2.5	1.624	1.882	1.55	2.282

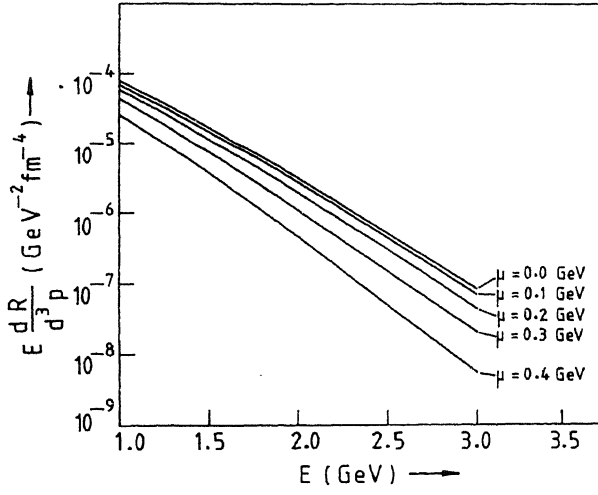


Figure 1. Thermal photon production spectrum $E\ dR/d^3p$ i energy density of the plasma ($\varepsilon = \text{GeV fm}^{-3}$) for various va

model equation of state

Acknowledgements

One of the authors (Pragya) is thankful to the Council of Scie
Research, India for financial support.

Note

After the completion of the paper our attention was drawn to a recent
(*Phys. Lett.* **B346**, 329 (1995)) who used the Bratten–Pisarski resum
baryon rich plasma and employed numerical methods to evaluate
integrals. Our results agree with them.

References

[1] D McLerran and T Toimela, *Phys. Rev.* **D31**, 545 (1985)
M Neubert, Z. *Phys.* **C42**, 231 (1989)
R Baier, H Nakkagawa, A Niegawa and K Redlich, Z. *Phys.* **C53**, 4
K Kajantie, J Kapusta, L McLerran and A Mekjian, *Phys. Rev.* **D34**,
J Clemans, J Fingberg and K Redlich *Phys. Rev.* **D35**, 2153 (1987)
M I Gorenstein and O A Mogilevsky, *Phys. Lett.* **B228**, 121 (1989)
[2] E V Shuryak, *Yad. Fiz.* **28**, 796 (1978) [*Sov. J. Nucl. Phys.* **28**, 408 (
K Kajantie and H I Miettinen, Z. *Phys.* **C9**, 341 (1981)
F Halzen and H C Liu, *Phys. Rev.* **D25**, 1842 (1982)
B Sinha, *Phys. Lett.* **B128**, 91 (1983)
R C Hwa and K Kajantie, *Phys. Rev.* **D32**, 1109 (1985)
[3] J Kapusta, P Lichard and D Seibert, *Phys. Rev.* **D44**, 2774 (1991)
[4] S Raha and B Sinha, *Phys. Rev. Lett.* **58**, 101 (1987)
S Chakrabarty, Jan-e Alam, D K Srivastava, B Sinha and Sibaji Raha
(1992)
[5] W Busza, *Nucl. Phys.* **A418**, 635C (1984); NA 36 Collaboration
E Anderson *et al, Nucl. Phys.* **A566**, 217C (1994); *Phys. Lett.* **B327**,
[6] H Sorge, H Stöcker and W Greiner, *Ann. Phys. (N.Y.)* **102**, 266 (198
R Mattielo, H Sorge, H Stöcker and W Greiner, *Phys. Rev. Lett.* **63**,
H Sorge, A V Keitz, R Mattiello, H Stöcker and W Greiner, *Phys. L*
A V Kreitz, L Winckelmann, A Jahans, H Sorge, H Stöcker and W G
353 (1991)

PRAMANA
—journal of
physics

© Printed in India

Some two body non-leptonic B decays and esti of f_D in the heavy quark effective theory

A K GIRI and R MOHANTA
Physics Department, Utkal University, Bhubaneswar 751 004, India

MS received 27 January 1997; revised 6 May 1997

Abstract. We have studied some of the two body nonleptonic decays of B m heavy quark effective theory using factorization hypothesis. Treating a_1 as have obtained its value as $a_1 = 0.822 \pm 0.045$, by correlating the experir branching ratio for $\bar{B}^0 \to D^+\pi^-$ process. With this value of a_1 the branchin other decay channels are in good agreement with the experiment. The decay c extracted to be 280.82 MeV. Heavy quark spin symmetry has also been teste

Keywords. Heavy quark effective theory; factorization approximation.

PACS Nos 11.30; 13.30

1. Introduction

Non-leptonic weak decays of hadrons contain a lot of valuab electroweak interaction of quarks. However these decays are nonpe and cannot be calculated reliably from the quantum chromo Lagrangian. A theoretical description of the nonleptonic decays of based on general principles is not yet possible. In fact some approxi used for these analyses. One of the techniques is the factorization hyp developed for weak hadronic decays of 'charm' and 'beauty' by Baue [1] and others [2], which considerably simplifies the necessary c

In this paper we analyse the class-I type decays i.e., $B \to D^+P^-$ $D^{*+}V^-$, where P and V stand for the pseudoscalar and vector mes Thus within the theoretical framework described here we are abl rates even if one of the decay products is not light. The use of H expression for the decay widths should contain the universal Isgur- is normalized at the point of zero recoil i.e., $\xi(v\cdot v' = 1) = 1$. Hc point of zero recoil, HQET does not predict the shape of the Isgu Hence to know the value of the Isgur–Wise function other than tl one has to rely on phenomenological models. Here we adopt the w Ali *et al* [8] to calculate the same.

The paper is organized as follows. In § 2 we present the gene study of non leptonic decays in the factorization method. The I evaluated in the wave function model [8] in § 3. Section 4 discussions.

2. General framework

Neglecting the penguin contribution, the effective Hamiltonian under consideration is given by

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{cb} V^*_{q_j q_i} [C_1(m_b)\mathbf{O}_1 + C_2(m_b)\mathbf{O}_2],$$

$$\mathbf{O}_1 = (\bar{q}_i q_j)^\mu (\bar{c}b)_\mu \quad \text{and} \quad \mathbf{O}_2 = (\bar{q}_i b)^\mu (\bar{c}q_j)_\mu,$$

where G_F is the Fermi coupling constant and the quark current $(\bar{q}_i$ (V-A) current. q_i and q_j are two types of quark flavors that are ha mesons. The values of the Wilson coefficients C_1 and C_2 are take

$$C_1(m_b) = 1.1 \quad \text{and} \quad C_2(m_b) = -0.24.$$

To evaluate the matrix elements of the effective Hamiltonian factorization hypothesis, where one assumes that the matrix ele operators may be expressed as the product of matrix elements Hence the amplitudes for $\bar{B}^0 \to D^+(D^{*+})M^-(M$ is either P or

where a_1 is related to C_1 and C_2 by

$$a_1 = C_1 + \frac{1}{N_c} C_2.$$

One expects N_c to be 3. However the phenomenology of charm decays
that $N_c \to \infty$ limit works well for D decays [1]. If this is also the cas
one would have $a_1 = 1.1$. On the other hand if $N_c = 3$ is used, one
Hence corrections of the order $1/N_c$ must be required. Therefore in thi
as a free parameter and determine it by comparing the predicted decay
with the experiment. Given the factorized Hamiltonian (4), one can th
amplitudes for the decays $\bar{B}^0 \to D^+ M^-$ and $\bar{B}^0 \to D^{*+} M^-$ as

$$\mathcal{M}(\bar{B}^0 \to D^+(D^{*+})M^-) = \frac{G_F}{\sqrt{2}} V_{cb} V^*_{q_j q_i} a_1 \langle M^-|(\bar{q}_i q_j)_\mu|0\rangle \langle D$$

To evaluate the factorized amplitudes we use the following matrix el

$$\langle P(p)|(\bar{q}_i q_j)_\mu|0\rangle = -i f_P p_\mu,$$
$$\langle V(p,\epsilon)|(\bar{q}_i q_j)_\mu|0\rangle = f_V \epsilon_\mu M_V,$$

where f_P and f_V are the pseudoscalar and vector meson decay cons
In the factorization approximation, two body nonleptonic decays
obtained by the semileptonic matrix elements of the weak currents
mesons. Therefore the matrix elements $\langle D^+(D^{*+})(v')|(\bar{c}b)_\mu|\bar{B}^0(v)\rangle$
their corresponding semileptonic decays. In the HQET these matrix
as [4]

$$\langle D^+(v')|\bar{c}b)_\mu|\bar{B}^0(v)\rangle = \sqrt{m_b m_c}\xi(w)(v+v')_\mu$$

and

$$\langle D^*(v',\epsilon)|(\bar{c}b)_\mu|\bar{B}^0(v)\rangle = \sqrt{m_b m_c}\,\xi(w)$$
$$\times (i\epsilon_{\mu\nu\alpha\beta}\epsilon^{*\nu} v'^\alpha v^\beta - (1+v\cdot v')\epsilon^*_\mu + (\epsilon^*\cdot v)v'_\mu),$$

Thus we obtain the partial decay widths as

$$\Gamma(\bar{B}^0 \to D^+P^-) = \frac{\tilde{G}_F^2}{8\pi}|\mathbf{p}|a_1^2 f_P^2 \xi^2(w) m_b m_c (1+w)^2 \Big(1 -$$

$$\Gamma(\bar{B}^0 \to D^+V^-) = \frac{\tilde{G}_F^2}{8\pi}|\mathbf{p}|a_1^2 f_V^2 \xi^2(w) m_b m_c (w^2-1) \Big(1 -$$

$$\Gamma(\bar{B}^0 \to D^{*+}P^-) = \frac{\tilde{G}_F^2}{8\pi}|\mathbf{p}|a_1^2 f_P^2 \xi^2(w) m_b m_c (w^2-1) \Big(1$$

$$\Gamma(\bar{B}^0 \to D^{*+}V^-) = \frac{\tilde{G}_F^2}{8\pi M_B^2}|\mathbf{p}|a_1^2 f_V^2 \xi^2(w) m_b m_c (1+w)$$

$$\times [(1+w)\{2M_V^2 + w(M_B - M_{D*})^2\} - (1-w)^2(M_B$$

where we have defined $\tilde{G}_F = (G_F/\sqrt{2}) V_{cb} V^*_{q_j q_i}$. The above expr widths (13–16) contain besides the known quantities, the unknown which can be calculated in the wave function model of Ali *et al* [

3. Evaluation of the Isgur–Wise function

The Isgur–Wise function which represents the nonperturbative QC of light cloud (spectator quark) rearrangement around the hea transition. Here we have used the wave function model of Ali *et al* The Isgur–Wise function calculated in this model predicts the $B \to K^*\gamma$ decays [8], the radiative decays of heavy mesons [10] an for nonleptonic decays of B mesons [7] very well. In this model the obtained from the overlap integral

$$\xi(v\cdot v') = \int r^2 \mathrm{d}r\, \phi_F^*(r)\phi_I(r) j_0(\Lambda r\sqrt{w^2-1}),$$

where ϕ_I and ϕ_F refer to the radial wave functions of the ini respectively. j_0 is the spherical Bessel function of the zeroth order. Λ is taken to be [11]

Table 1. Predictions for the branching ratios ($\mathcal{B}$) for the nonleptonic

Decay process	$\lvert\mathbf{p}\rvert$ in MeV	$\xi(v\cdot v')$	$\mathcal{B}$ (Present calculation)	Exp
$\bar{B}^0 \to D^+\pi^-$	2306.278	0.7632	3.0×10^{-3}	(3.0 ±
$\bar{B}^0 \to D^+K^-$	2278.875	0.7681	2.296×10^{-4}	
$\bar{B}^0 \to D^+D^-$	1863.654	0.8382	6.0×10^{-4}	6.0
$\bar{B}^0 \to D^+D_s^-$	1812.223	0.8463	7.794×10^{-3}	(7 ±
$\bar{B}^0 \to D^+\rho^-$	2236.221	0.7756	7.283×10^{-3}	(7.8 ±
$\bar{B}^0 \to D^+K^{*-}$	2210.937	0.7801	3.959×10^{-4}	
$\bar{B}^0 \to D^+D^{*-}$	1789.684	0.8498	4.1×10^{-4}	
$\bar{B}^0 \to D^+D_s^{*-}$	1731.368	0.8587	5.048×10^{-3}	(2.0 ±
$\bar{B}^0 \to D^{*+}\pi^-$	2254.486	0.7724	2.578×10^{-3}	(2.6 ±
$\bar{B}^0 \to D^{*+}K^-$	2225.952	0.7774	1.939×10^{-4}	
$\bar{B}^0 \to D^{*+}D^-$	1789.684	0.8498	3.688×10^{-4}	
$\bar{B}^0 \to D^{*+}D_s^-$	1734.989	0.8582	4.563×10^{-3}	(1.2 ±
$\bar{B}^0 \to D^{*+}\rho^-$	2181.486	0.7852	7.382×10^{-3}	(7.3 ±
$\bar{B}^0 \to D^{*+}K^-$	2155.096	0.7898	4.233×10^{-4}	
$\bar{B}^0 \to D^{*+}D^{*-}$	1710.961	0.8618	9.148×10^{-4}	
$\bar{B}^0 \to D^{*+}D_s^{*-}$	1648.612	0.8710	1.231×10^{-2}	(1.9 ±

normalization condition $\xi(1) = 1$. Thus we obtain $\xi(w)$ to be

$$\xi(w) = \exp(-a^2/4\beta_{\text{ave}}^2),$$

where $a = \Lambda(\sqrt{w^2 - 1})$ and $\beta_{\text{ave}} = (\beta_B + \beta_D)/2$.

The value of w (i.e., $v\cdot v'$) is determined by considering the kinematics o
Since we are dealing with the two body decays $\bar{B}(v) \to D(v')P(p)/V(p)$
conservation implies

$$M_B v^\mu = M_D v'^\mu + p^\mu,$$

which gives

$$v\cdot v' = \frac{M_B^2 + M_D^2 - M_{P/V}^2}$$

width for the process $\bar{B}^0 \to D^+\pi^-$ with the corresponding exper have obtained the value of a_1 as

$$|a_1| = 0.822 \pm 0.045,$$

which is in general agreement with the result predicted by Du an analysis Gourdin *et al* [15] have determined the value of a_1 from hadronic decays. Assuming factorization, they have used the hac $B \to D(D^*)$ transitions from various models, whereas for experi taken as the world average data and obtained the ranges for a_i i

$$0.91 \leq a_1 \leq 0.97 \text{ BSW}$$
$$0.96 \leq a_1 \leq 1.03 \text{ HQET I}$$
$$0.96 \leq a_1 \leq 1.07 \text{ HQET II (One standard deviation)}$$
$$0.87 \leq a_1 \leq 1.17 \text{ HQET II (Two standard deviation)}$$

Our estimated value of a_1 is lying below these values. With this p obtain the branching ratios for other nonleptonic $\bar{B}^0$ decay chann table 1. Comparing the experimental result $\mathcal{B}(\bar{B}^0 \to D^+D^-) =$ predicted value, the decay constant of D meson is extracted to b

$$f_D = 280.28\,\text{MeV},$$

which is well below the experimental upper limit $f_D < 310\,\text{MeV}$ [spin symmetry implies

$$\mathcal{B}(\bar{B}^0 \to D^+\pi^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-) = 1$$

and

$$\mathcal{B}(\bar{B}^0 \to D^+\rho^-)/\mathcal{B}(\bar{B}^0 \to D^{*+}\rho^-) = 1.$$

Since the quark masses are non-infinite, these expressions are unity. Therefore as a test of heavy quark spin symmetry we h predicted results

$$\mathcal{B}(\bar{B}^0 \to D^+\pi^-) = 1.16\,\mathcal{B}(\bar{B}^0 \to D^{*+}\pi^-)$$

meson easily escaping the region of interaction with other quarks, there effect of the final state interactions (FSI) [18]. In recent past there ha interest [19–22] in exploring the role played by non-factorized term decays of 'charm' and 'bottom' mesons. In ref. [19] the authors ha calculate the nonfactorized contributions to two body hadronic decays been realized recently that, within the factorization approximation all th models for $B \to \psi K(K^*)$ transitions fail to reproduce the measured valu polarization fraction $(\Gamma_L/\Gamma)(B \to \psi K^*)$ and the production rat $\Gamma(B \to \psi K)$. These decays being color suppressed (amplitude propor a smaller branching fraction than the color favored modes ($B^0 \to D^+$ amplitudes are proportional to a_1. Thus a nonfactorizable effect c contribute a larger fraction of the total amplitude to the color suppresse color favored ones [21]. This fact is also pointed out in ref. [2 factorization hypothesis works better than the N_c assumption for colo because nonfactorizable effects are small in these decays. Therefore assume that the color favored two body non-leptonic weak decays factorizable contributions whereas nonfactorizable amplitudes play a understanding color suppressed $B \to \psi K(K^*)$ decays.

In summary, we have performed an analysis of two body non-lept meson in the factorization approximation. The matrix elements of the evaluated using the heavy quark effective theory. The universal Isgu calculated using the wave function model of Ali *et al* [8]. Although i Liu have analysed the nonleptonic B decays using heavy qua factorization approximation, they have used the QCD sum rule appro Wise function whereas in our case we have taken the wave functi obtained the result in a straight forward manner. Using this value of a_1 the branching ratios for other decays, which are in good agreement data. The decay constant of D meson is extracted to be $f_D = 280.28$ Me working approximation adopted here the present calculation provides a to explain reasonably the two body non-leptonic decays of $\bar{B}^0$ meson

Acknowledgements

[3] N Isgur and M B Wise, *Phys. Lett.* **B232**, 113 (1989); *Phys. Lett.* *Phys.* **B348**, 276 (1991)
[4] A F Falk, H Georgi, B Grinstein and M B Wise, *Nucl. Phys.* **B343**
[5] H Georgi, *Phys. Lett.* **B240**, 447 (1990); *Nucl. Phys.* **B348**, 293 (1
[6] M Neubert, *Phys. Rep.* **245**, 259 (1994)
[7] A K Giri, L Maharana and R Mohanta *Pramana – J. Phys.* **48**, 10
[8] A Ali, T Ohl and T Mannel, *Phys. Lett.* **B298**, 195 (1993)
[9] G Altarelli, G Curci, G Martinelli and S Petrarca, *Nucl. Phys.* **B18**
[10] A K Giri, L Maharana and R Mohanta, *Pramana – J. Phys.* **47**, 22
[11] T Altomari, *Phys. Rev.* **D37**, 677 (1988)
[12] N Isgur, D Scora, B Grinstein and M B Wise, *Phys. Rev.* **D39**, 79
[13] Particle Data Group; Review of Particle Properties, *Phys. Rev.* **D5**
[14] D Du and C Liu, *Phys. Rev.* **D48**, 3397 (1993)
[15] M Gourdin, A N Kamal, Y Y Keum and X Y Pham, *Phys. Lett.* **B**
[16] P McBride, *Nucl. Phys.* (*Proc. Suppl.*) **B50**, 240 (1996)
[17] W Ross, *Nucl. Phys.* (*Proc. Suppl.*) **B39**, 381 (1995)
[18] A N Kamal, Q P Xu and A Czarnecki, *Phys. Rev.* **D49**, 1330 (199
[19] B Block and M Shifman, *Nucl. Phys.* **B389**, 534 (1993)
M Shifman, *Nucl. Phys.* **B388**, 346 (1992)
[20] M Gourdin, A N Kamal and X Y Pham, *Phys. Rev. Lett.* **73**, 3355
A N Kamal and A B Santra, Z. *Phys.* **C72**, 91 (1996)
[21] A N Kamal and A B Santra, *Phys. Rev.* **D51**, 1415 (1995)
[22] H Y Cheng, Z. *Phys.* **C69**, 647 (1996)

PRAMANA
—journal of
physics

© Printed in India

Modified WKB transmission for fusion

S V S SASTRY and S K KATARIA
Nuclear Physics Division, Bhabha Atomic Research Centre, Mumbai 400 0

Abstract. In the optical model (OM) approach for fusion, absorption of fl barrier position is presented in detail at low energies. It has been shown th can be well approximated as a sum of the WKB transmission and a long r contribution. Owing to absence of LRA, the fusion predictions of coupled transmission approach like the CCFUS code, do not agree with the predictio reaction channel (CRC) calculations based on OM approach using the code code with a modified transmission which includes LRA contribution is shov the CRC results using FRESCO. The static deformation of the colliding nu the fusion imaginary potential and therefore the deep sub-barrier fusion cr

Keywords. Sub-barrier fusion; CRC calculations; barrier penetration; WK model; long ranged absorption; deformed potentials.

PACS Nos 25.70; 24.10

1. Introduction

The study of heavy ion (HI) reactions over the past decade has yielc about the reaction mechanism and the inter-relation between vario The coupled channels (CC, also denoted by CRC) formalism i explaining most of the experimental data of HI fusion around the ba the channel coupling effects on elastic and various other reactio fusion, as described by the CC formalism, is estimated as the flux the coupled channels system by using a short ranged imaginary pote

estimation of coupling effects [5] (only on fusion channel). The
Hill–Wheeler transmission formula in addition to various ap
diagonalize the coupled equations [5]. These approximations grea
channel problem and thus reduce the computation time and memo
at the cost of loosing information about other channels. Conseque
successfully used to understand the experimental fusion data
around the barrier energies.

However, the recent extensive study of the sub-barrier fusio
energies and the experimental measurements of fusion mean s
eq. (1a)] by various methods have led to new results that are no
current model predictions [1, 2]. These results can be summariz
enhancement of FMSS as derived from the fission angular anisot
(ii) the anomalous enhancement of fusion cross section at deep s
The CCFUS code fails to qualitatively explain these anomalo
demonstrated by Satchler *et al* that a long ranged optical potentia
the anomalous FMSS values [7]. In this method, the flux is absor
contrast to the requirement of complete traversal of barrier of the
Thus it is necessary to examine these approaches for fusion i
believed that the barrier penetration model and the optical mod
ranged imaginary potential ($r_i = 1.0$ fm) give similar results for
[8]. However, it was shown for HI systems at deep sub-barr
transmission method underpredicts the FMSS of an OM calcula
predictions of coupled channel codes based on WKB transmissio
CCFUS) will not be in agreement with FRESCO calculations
reported in ref. [9]. In § 2, we present the radial distribution flu
different energies and a simple correction to barrier transmission
OM absorption of flux. In § 3, a simple procedure will be pres
long range absorption (LRA) effects in the barrier transmission fo
couplings. In § 4, the present work will be applied to analyse the
of some recently reported heavy ion systems. In §5, we present th
systems with static deformation and in § 6, a brief summary.

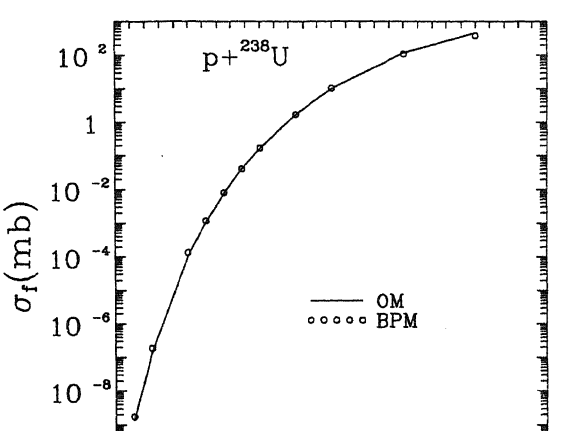


Figure 1. Comparison of the OM absorption excitation
excitation function from transmission method for the case of
represents OM results and the circles represent the transmi

that the OM absorption cross sections agree well with the results o
flux through the corresponding real barrier. This agreement of res
barrier energy of 2 MeV of proton has been observed in spite of us
($r_i = 1.355$) and large diffuseness (0.635 fm) for this system in con
a heavy ion case shown in ref. [9]. This happens because
corresponding R_m values [cf. eq. (4)] for light ion case, where t
turns out to be very large. Therefore, the LRA effect is insignifica
as the $W(r)$ in this region is extremely small, unlike the case for

In the following, we present the results for heavy ion syster
absorption of flux in the OM depends on the choice of imaginary
extends over the entire domain of the radial integration. The OM
flux absorbed ($A_l(r)$) and the cross section (σ_l) are related by

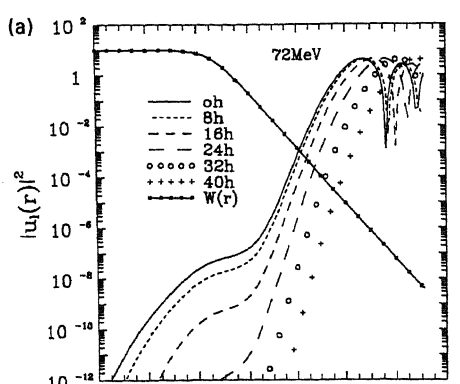


Figure 2. (a) Radial density of the wave function, $|u_l(r)|^2$ energy for different partial waves as labelled in the figure. $W(r)$, is shown by dashed curve. The OM code SNOOPY i (Woods–Saxon form) parameters as follows: $r_c = 1$ $r_0 = 1.179$ fm, $a_0 = 0.685$ fm, $W = 10$ MeV, $r_i = 1.0$ fm, distribution of absorption of flux $(A_l(r))$ in OM corresp different partial waves shown by different symbols. The values obtained from eq. (4) are indicated in parentheses.

significant contribution beyond the barrier position at very low $W(r)$ and the l value. The wave function exhibits strong oscilla which are weighted by the $W(r)$. Therefore, $A_l(r)$ exhibits multip one is around the barrier position ($R_b = 11$ fm) as used in transmission models. The next significant peak in $A_l(r)$ occur turning point (R_0) and corresponds to absorption beyond the becomes obvious when eq. (1) is written as

$$\sigma_l = \frac{-8\,\pi}{\hbar v\, k^2}(2l+1)\left[\int_0^{R_b} |u_l(r)|^2 W(r)\mathrm{d}r + \int_{R_b}^{\infty} |u_l(r)\right.$$

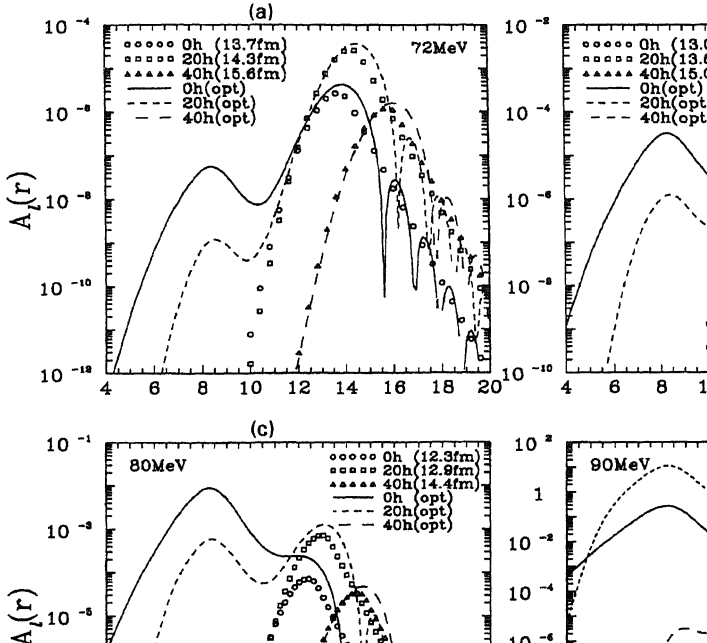
with $k(r) = k((2\eta/kr) + (L^2/k^2r^2) - 1)^{1/2}$, η is the Sommerfel
the incident wave number. For radial separations beyond $R_0, u_l(r$
Though this is overestimation of the flux at large distances, it is
Under these approximations, the second part of eq. (2a) can also

$$\sigma_l^{\mathrm{LRA}} = -C_l \int_{R_b}^{\infty} \mathrm{d}r \exp\left(-\kappa r + 2\int_{R_0}^{r} k(r')\,\mathrm{d}r'\right)$$

with

$$C_l = 10(2l+1)\pi/k^2 8/\hbar v W_0 \mathrm{e}^{\kappa r_i}$$

and $k(r')$ is the local wave number in the presence of Coulomb an
and $\kappa = 1/a_i$ is the inverse of diffuseness of the fusion imagin



eq. (1). Following the procedure of ref. [10] (similar integrals for distance, R_m, where the LRA contribution is maximum can be obt

$$R_m = \frac{\eta}{kq^2}\left(1+\left(1+\frac{q^2L^2}{\eta^2}\right)^{1/2}\right) \quad \text{and} \quad q^2 = 1+\kappa^2/$$

The integrand of eq. (3b) can be approximated as a gaussian cent ref. [10]

$$\sigma_l^{\mathrm{LRA}} = 2C_l\int_{R_0}^{R_m} \mathrm{d}r \exp\left(-\kappa r + 2\int_{R_0}^{r} k(r')\mathrm{d}r'\right).$$

One can also use eq. (3b) replacing R_b and ∞ of outermost integr limits beyond the barrier position. In the present work, we used e taken are 12 fm and 20 fm respectively. Further, for integratio trajectory, the integrand of eq. (3a) and also the corresponding te replaced by an analytic expression.

This integrand of eq. (3b) (alongwith multiplicative factor $-C_l$) symbols in figures 3(a, b, c, d). As seen in figures, this approximatio $A_l(r)$ obtained from OM code, represented by various continuous c represented by eq. (4) for different partial waves are indicated in pare and also in figures 3(a, b, c, d). As seen in figure, at high energies th R_b only contributes. This is the usual WKB transmission of flux thro and is independent of W_0 as represented by the first term of eq. (2b) that the second term of eq. (2b) contributes significantly, which con of the OM approach and is responsible for large OM estimates of FM if the second term is neglected, the OM also predicts saturation valu WKB method. Thus the LRA effects will be seen in any CRC c approach for fusion.

3. Coupled channels method for fusion

In the CC formalism, there are several channels from which absor

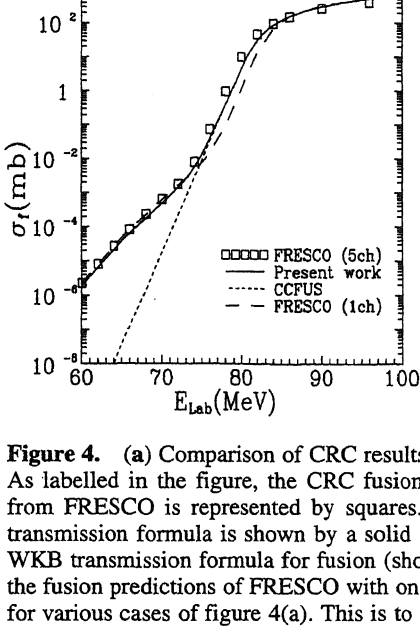


Figure 4. (**a**) Comparison of CRC results and the results of
As labelled in the figure, the CRC fusion (absorption) exci
from FRESCO is represented by squares. The results from
transmission formula is shown by a solid curve along with t
WKB transmission formula for fusion (short dashed curve).
the fusion predictions of FRESCO with only elastic channel.
for various cases of figure 4(a). This is to compare the modi
CRC calculations.

where α and β are the various channels included in the CRC calcu
the outgoing wave function for relative motion of coupled equatic

It should be noted that the term fusion here refers to removal of
potentials from the coupled channels system. Therefore, the fusi
in these calculations refers to absorption excitation function.
function by the CRC method for different a_i values was studied
seen that as long as the fusion potential is very short ranged,
fusion gives the same results as WKB transmission method.
parameter of fusion imaginary potential serves as a measure o
OM approach. The CRC fusion excitation function differs fro

As discussed earlier, the CCFUS code based on WKB transmissi equivalent to complete CRC calculations using the codes like Figure 4(a) shows the fusion excitation function obtained from FR and from the CCFUS code with WKB transmission approach (sho anomalous enhancement of the CRC results (squares) compared to curves) results is owing to LRA effect. Figure 4(b) shows the results. In order to remove this defficiency, the WKB transmissio modified by using only elastic channel represented by $\sigma_l^{(m)}$,

$$\sigma_l^{(m)} = \frac{\pi}{k^2}(2l+1)\,T_l^{\mathrm{CC}} + \sigma_l^{\mathrm{LRA}}.$$

Here, T_l^{CC} is the transmission for fusion from usual unmodifie includes all the effects of channel couplings. The second term σ_l^{LRA} In addition, the WKB radial wave function part in eq. (3b) was expression for a Coulomb trajectory. Under these conditions, the i be seen to be very simple and is evaluated once for each partial wa number of channels coupled. Therefore, the computation time and are almost unchanged.

Figure 4(a) also shows the fusion excitation function by the mo using eq. (6d) (solid curve), FRESCO (squares) and the unmodified dashes) for a_i value of 0.4 fm (equivalent to 0.8 fm in FRESCO w The W_0 used is 10 MeV and r_i is 1.0 fm which are same as CRC in the corresponding FMSS values. As shown in figure 4(a,b), th CCFUS code compare well with CRC results using FRESCO. between these two cases means that the proposed modification is for simulating the results of realistic CRC codes. Therefore, this m for parametrizing the experimental data on heavy ion fusion.

4. Deep sub-barrier fusion of heavy ion systems

Using the modified CCFUS code, we have calculated the fusion of systems. The imaginary potential parameters used in CC

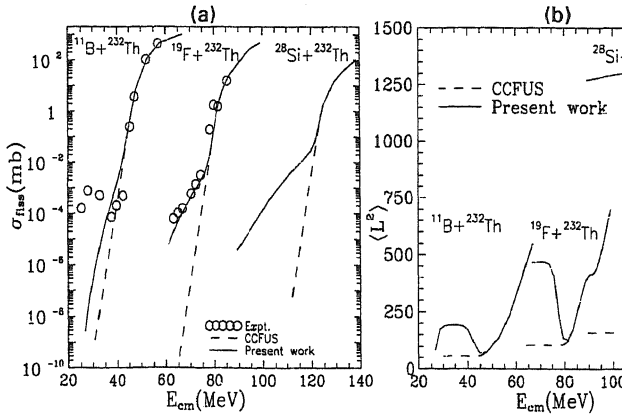


Figure 5. (**a**) Comparison of the deep sub-barrier fusion prec method for different systems with same target. The results from (transmission formula are shown in solid curve along with the WKB transmission formula for fusion (dashed curve). (**b**) FM various cases of figure 5(a). This is to compare the results o modified and WKB transmission formulae.

It should be noted that the imaginary potential parameters are l a comprehensive coupling scheme with all significantly contril considered for CC calculations. Consequently, the enhancem function and mean square spin arising due to LRA effects, which de potential, would also change. Therefore, only the qualitative model predictions is of importance. In addition, it should be nc enhacement observed in deep sub-barrier fusion cross section has the LRA effects alone, resulting from spherical imaginary potentials mean square spin values predicted would provide an upper limit to th for these systems. The effect of static deformation of the potentials o be investigated and will be discussed in the next section.

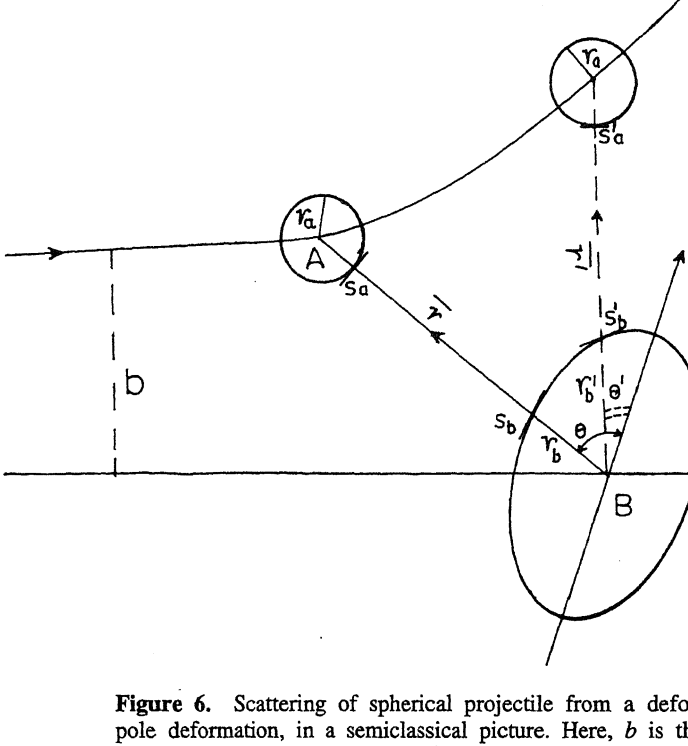


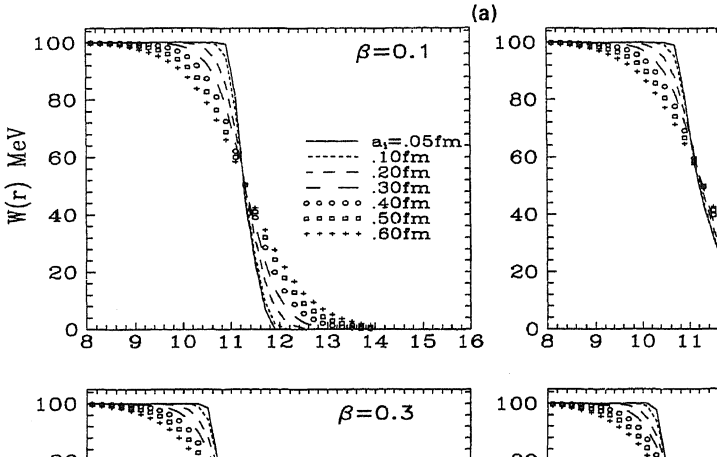
Figure 6. Scattering of spherical projectile from a defo
pole deformation, in a semiclassical picture. Here, b is th
is the projectile radius, r and r' are the center to center
ions at the positions P and Q along the trajectory. r_b and
deformed target as seen by the projectile at P and Q res
tion dependent imaginary potential uses the surface to
the radial direction at P and Q which are given by $r - ra$ -

between centers of the nuclei is given by

$$W(r,\theta) = \frac{-W_0}{1+\exp((r-r_0(\theta))/a_0)}.$$

The orientation dependence of the potential half radius of eq. (7c) is a
for the case of scattering from the deformed target with only quadrup
and r'_b are the radii of the deformed target as seen by the projectile a
trajectory. The deformation dependent imaginary potential uses the
distances along the radial direction at P and Q which are given
$r' - ra - r'_b$, of eq. (8)

Figure 7(a), having four figure sets, shows the effective W as
different deformation parameter values. Each of these four figures sh
effects for various diffuseness values of the potential. All the potenti
volume Woods–Saxon type and the spherical potential radius para
1.2 fm. As can be seen from the figure, the effective potential behave
larger diffuseness parameter in the neighbourhood of the mean radius
it decreases with the same diffuseness as that of the spherical potenti



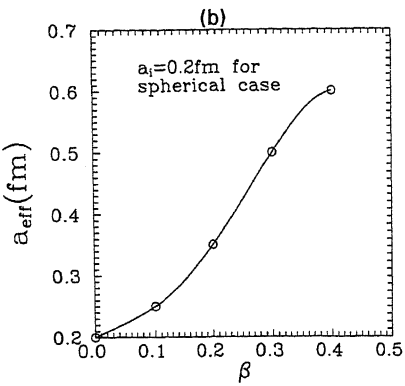


Figure 7(b). Effective diffuseness as a function of deform radius parameter is chosen as 1.2 fm.

effect is more pronounced for potentials of smaller diffuseness and potentials of very large diffuseness. For the curve with 0.2 deformation parameter values, the effective spherical Woods–Saxc the region of 11 fm to 13.5 fm has a diffuseness of 0.5 fm. F deformation dependence of effective diffuseness of a Woods–Sa region of 11 fm to 14 fm, which has 0.2 fm diffuseness for co potential. This is the case for which the deformation effect is seen the effect of deformation is that the diffuseness parameter cha before deformation is switched on) to an effective value of 0.5 fm. from the figure that the curves representing the different cases of { similar in magnitude. Therefore, the effective potential depends both deformation and spherical diffuseness. Since, the fusion imag approach are of shorter range and smaller diffuseness, the deform extend the imaginary potential beyond the barrier radius even th potential is very short ranged. Consequently, this effect can absorption of flux for higher partial waves and therefore results fusion mean square spin.

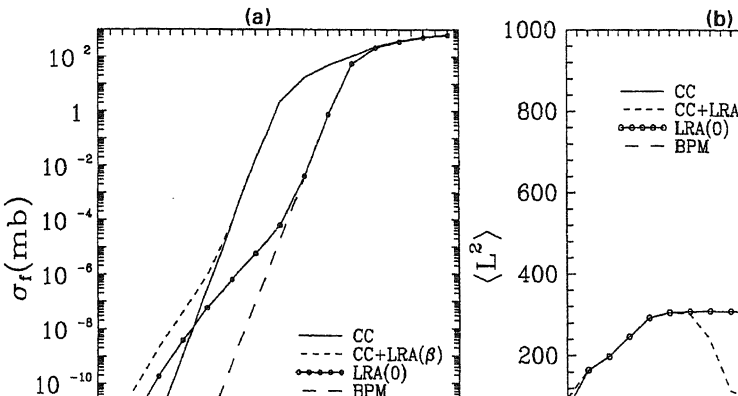


Figure 8(a, b). Effect of deformation of W on LRA contribu system. Figure 8(a) represents the excitation function and figure ing mean square spin values. The parameters are $\beta_2 = 0.3$, $r_0 = 1$ is 0.3 fm. The long dashed curve represents the one dimension model results. The solid curve labelled CC is the usual CCFUS Wong's method in the presence of static deformation. The open by solid curve) represent the OM and LRA effects, but wit imaginary potential. The short dashed curve represents the OM in the presence of deformed imaginary potential.

imaginary potential i.e. without deformation effects in W, as represent interpolated by solid curve. This is equivalent to simple optical mod spherical imaginary potential. The third case is for absorption incl presence of deformation of W as shown by short dashes. Compari shows that the minimum in the FMSS value shifts to lower energie compared to squares) and the peak value of the FMSS enhacement at s remains the same. At very low energies, these results merge with

optical model LRA effects, deformation dependence of imaginary inclusion of nonlocal effects is being studied [12].

6. Summary

It is shown that in the OM, absorption beyond the barrier po contribution to fusion at very low energies. This OM results can be WKB transmission and an LRA part estimated using WKB elas CRC calculations show that the absorption (mainly LRA) into fusi is the dominant contribution at deep sub-barrier energies. The resu based on transmission approach, which is very commonly used are not in agreement with the CRC results at low energies. The fusion in CCFUS code has been modified incorporating the agreement with FRESCO code, without much additional computa requirements. The static deformation effect on the short ranged in been studied using the effective potential as a sum over all the ori potential becomes more diffused in the vicinity of mean radius a diffuseness at large distances. The LRA effects corresponding deformed imaginary potential, summed over all orientations, shi compared to undeformed case.

Acknowledgements

The authors are thankful to Drs S S Kapoor, A K Moha D M Nadkarni and A Saxena for many discussions during the c acknowledge the fruitful collaboration of Prof. I J Thompson du work as well as for providing us the FRESCO code.

References

[1] R Vandenbosch, *Annu. Rev. Nucl. Part. Sci.* **42**, 447 (1992)
[2] G R Satchler, *Phys. Rep.* **199**, 147 (1991)
[3] I J Thompson, *Comput. Phys. Rep.* **7**, 167 (1988)

PRAMANA
—journal of
physics

© Printed in India

Excitation functions of alpha particle induced aluminium and copper

S MUKHERJEE*, B BINDU KUMAR and N L SINGH†
School of Studies in Physics, Vikram University, Ujjain 456 010 India
† Physics Department, M.S. University, Baroda, India
* Corresponding author.

MS received 6 August 1996; revised 18 March 1997

Abstract. Cross-sections for the reactions with product nuclei ^{24}Na, ^{22}Na investigated over the energy range of 30 to 75 MeV for alpha particle induc aluminium and copper, using stacked-foil activation technique. The measur were analysed with special reference to their suitability for monitoring bear The experimental results were compared with the predictions of hybrid assumption of initial exciton number $n_0 = 4(4p0h)$ best satisfies the measur in the present work.

Keywords. Nuclear reactions; stacked foil activation technique; hybrid m

PACS No. 25.60

1. Introduction

Nuclear reactions are frequently understood in terms of two extreme of their time scale of occurence. These are the fast direct reacti compound nucleus reaction mechanisms [2].

However, in recent years, there has been an increasing evidence types of processes, known as pre-equilibrium reaction mechanism

nuclear reaction studies with low and intermediate energy alpha p
the conditions of ^{24}Na, ^{22}Na, ^{68}Ga and ^{67}Ga (targetting, yields, p
of the excitation function was necessary. The alpha particle
aluminium were earlier measured by a few groups [15–1
disagreement among themselves, with some measurements havi
20–40% [15]. Even though large number of experimental data ar
[18–25] for the alpha induced reactions producing ^{68}Ga and ^{67}G
obtained using poor resolution detectors. However, some of the r
been done using good resolution Ge(Li) detectors [22–25].

In view of the large uncertainties and mutual discrepancies
measurement of these reactions are still needed. With the advent
large volume semiconductor detectors and with the availability of
excitation functions measurement has entered into a new era. The
is carried out with the following motivations:

(1) determination and optimisation of the ^{24}Na, ^{22}Na, ^{68}Ga and
induced reactions on aluminium and copper,
(2) to compare the experimentally obtained data with the theoreti
[7, 26, 27],
(3) to study the reaction mechanism involved in the production c
^{67}Ga isotopes.

2. Experimental procedure

2.1 *Irradiation and counting*

Stacked-foil activation technique was employed in the presen
supporting metallic foils of aluminium and copper of purity grea
thicknesses 14 mg/cm^2 and 23 mg/cm^2, respectively were us
irradiation. Copper foils were interspersed with aluminium foil
acting as energy degraders, while the aluminium stack was prepar
foils. The irradiation of the stacks were carried out at the Vari
Centre, Calcutta. The primary beam of the alpha particles wa

Table 1. Reaction studied in the present work and the relevan of residual nuclei.

Reaction	Residual nucleus	Q-Value (MeV)	Half-life $T_{1/2}$
$^{27}Al(\alpha, \alpha 2pn)$	^{24}Na	−31.40	15.05 *h*
$^{27}Al(\alpha, 2\alpha n)$	^{22}Na	−22.50	2.60 *Y*
$^{65}Cu(\alpha, n)$	^{68}Ga	−5.85	68.00 *m*
$^{65}Cu(\alpha, 2n)$	^{67}Ga	−14.10	28.26 *h*

insulated irradiation head serving as a kind of Faraday cup [17–2 electrons were prevented from escaping. Using this charge, the flux

2.3 *Cross-section determination*

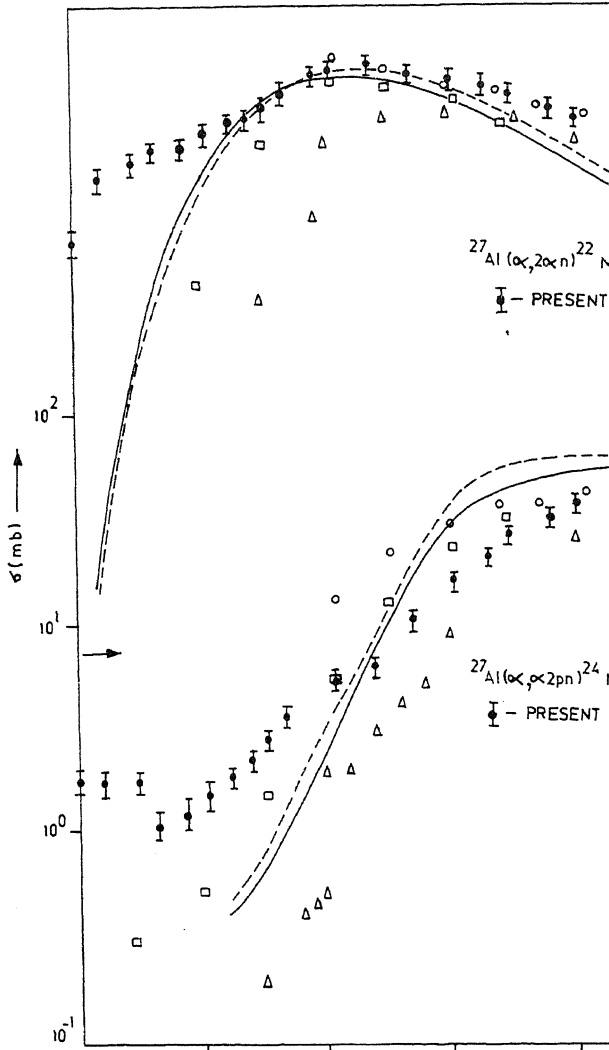
The formula used for the determination of reaction cross section is

$$\sigma = \frac{A_\gamma A_{gm} \lambda}{\phi W_i P_i P_\gamma \theta_\gamma N_{Av} [1 - \exp(-\lambda t_i)][\exp -\lambda t_w][1 - \exp(-}$$

where, σ is the cross section for the reaction, A_γ is the photopeak area gamma ray of the residual nucleus, A_{gm} is the gram atomic weight o λ is the disintegration constant of the residual nucleus, ϕ is the particles, W_i is the weight per unit area of the target foil, P_i is the frac weight of the target isotope of interest, θ_γ is the fraction of the chara emitted, P_γ is the photopeak efficiency of the gamma rays, N_{Av} is th t_i, t_w, Δ are the periods of irradiation, waiting and counting respecti

2.4 *Errors*

The total error in the experimental cross section is contributed by photopeak areas of characteristic gamma-ray, 1–8% error in their a


^{27}Al (α, 2αn)^{22}
– PRESENT
^{27}Al(α, α2pn)^{24}
– PRESENT
σ (mb)
10^2
10^1
10^0
10^{-1}

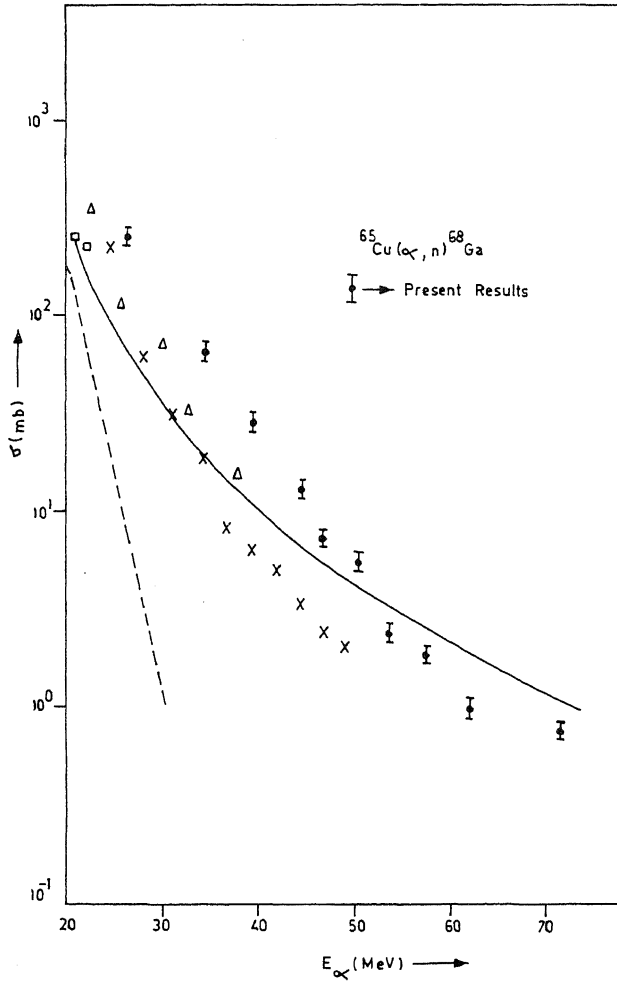
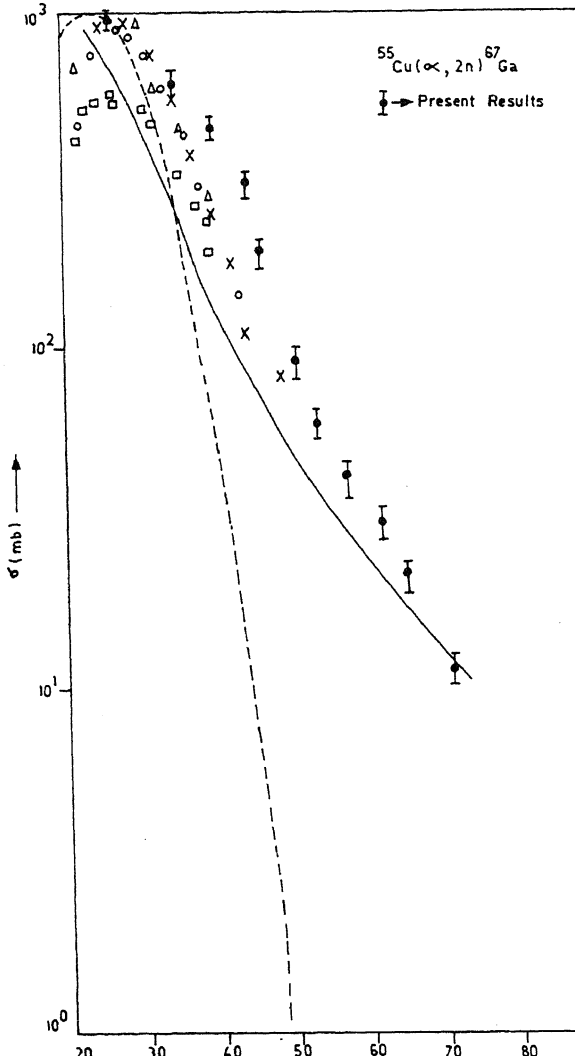


Figure 2. Comparison of the present and previous experime
the theoretical excitation functions of $^{65}Cu(\alpha, n)^{68}Ga$ reaction


$^{55}Cu(\alpha, 2n)^{67}Ga$
Present Results
10^3
10^2
10^1
10^0
σ (mb)
20
30
40
50
60
70
80

to 75 MeV. This can help avoid errors in the flux determination the improper estimation of the middle energy of the bombardir monitor foil. The excitation functions for the other two product nuc show a systematic trend in the entire region starting from 30 to 7 these reactions may also be used for the determination of flux of the particle.

4. Comparison with hybrid model

In the present work the excitation functions are calculated on the b model [7, 26, 27] with and without the inclusion of pre-equilbrium em Basically the hybrid model is semi-classical and hence involves a physical parameters along with a few adjustable parameters. A short options chosen is given below. The nuclear masses were calculated Swiatecki mass formula [29] with the liquid drop having shell corre pairing, i.e. the level density pairing absorbed in binding energies. sections were calculated by optical model subroutine included in the co parameter of $A\backslash 8\,\mathrm{MeV}^{-1}$ was used. For the equilibrium part, the approximation is used which takes into account the full gamma, compe emission. The initial exciton number $n_0(P_0h_0)$ plays an important role the entire cascading process of binary collisions and thereby influence component in the particle spectra.

We have made the theoretical calculations using the initial tions $4(4p0h)$, $5(5p0h)$ and $6(5p1h)$. It was observed that the theo of the hybrid model with $4(4p0h)$ gives better results, as compared of the $5(5p0h)$ and $6(5p1h)$ configurations with respect to initial co and Mignerey [30] used $n_0 = 4(4p0h)$ to calculate (α, p) spectra *et al* [31] have also discussed this point in detail and recomm application of $n_0 = 4$. Therefore we have shown $4(4p0h)$ con comparison. Evidently, an initial exciton $n_0 = 4(4p0h)$ configuration to a break up of the incoming alpha particle in the field of the nucleons occupying excited states above the Fermi level gives a bette

energy range 40 to 80 MeV is by statistical evaporation process. were also made by Bowman and Blann in their study of mean r induced reactions on ^{27}Al. They found that there is only 20 to 2 transfer for ^{24}Na and ^{22}Na, confirming compound nucleus mechan the production of ^{24}Na and ^{22}Na.

Figures 2 and 3 shows the reactions ^{65}Cu(α, n) and ^{65}Cu$(\alpha, 2$ of ^{68}Ga and ^{67}Ga respectively. In (α, n) reaction, as shown in fi nucleus mechanism, given by Weisskopf–Ewing estimate, is do part of the excitation function, while there are considerable p beyond 30 MeV of bombarding energies. Similar arguments can al reaction, where the pre-equilibrium effects predominate beyond energy.

It can be seen in figure 2 that there is a clear shift in energy betw experimental compound nucleus peaks. Generally, such shifts complete neglect of angular momentum effects in the Weissk calculations of the compound nucleus contributions. More elab Hauser–Feshbach theory may bring about a better agreement.

5. Conclusions

Four excitation functions for the production of ^{22}Na, ^{24}Na, ^{67}Ga a from the alpha bombardment of aluminium and copper were meas up to 75 MeV. The high resolution Ge(Li) detectors were em improvement in the quality of data was observed. In all the four r ^{24}Na, ^{67}Ga and ^{68}Ga, the excitation functions for the reaction pr were studied for the first time in the energy range 50 to 75 MeV.

A comparison of the experimental data with the results of hyb were performed with and without using the pre-equilibrium emis code. The initial exciton configuration was varied betw $n_0 = 6(5p1h)$. However, $n_0 = 4(4p0h)$ gave the best agreement b experiment. A pronounced pre-eqilibrium emission of neutron reactions with product nuclei ^{67}Ga and ^{68}Ga beyond 30 and However the reactions with product nuclei ^{24}Na and ^{22}Na ma

References

[1] T Tamura, *Ann. Rev. Nucl. Sci.* **19**, 99 (1969)
[2] A Fleury and J M Alaxender, *Ann. Rev. Nucl. Sci.* **24**, 279 (1974)
[3] M Blann, *Ann. Rev. Nucl. Sci.* **25**, 123 (1975)
[4] J J Griffin, *Phys. Rev. Lett.* **17**, 478 (1966)
[5] G D Harp and J M Miller, *Phys. Rev.* **C3**, 1847 (1971)
[6] E Gadioli, E Gadioli-Erba and P G Sona, *Nucl. Phys.* **A217**, 589 (1973
[7] M Blann, *Phys. Rev. Lett.* **27**, 337 (1971)
[8] C K Cline and M Blann, *Nucl. Phys.* **A172**, 225 (1971)
[9] C Kalbach, *Phys. Rev.* **C23**, 124 (1981)
[10] D Agassi, H A Weiden Muller and G Mantzouranis, *Phys. Rev.* **22**, 145
[11] T Tamura, T Udagawa, D H Feng and K K Kan, *Phys. Lett.* **B68**, 109
[12] L Avaldi, R Bonetti and L Colli-Milazzo, *Phys. Lett.* **B94**, 463 (1980)
[13] H Feshbach, A Kerman and S Koonin, *Ann. Phys.* **125**, 429 (1980)
[14] G M Field, R Bonetti and P E Hodgson, *J. Phys.* **G12**, 93 (1986)
[15] W W Bowman and M Blann, *Nucl. Phys.* **A131**, 513 (1969)
[16] H J Probst S M Qaim and R Weinreich. *Int. J. Appl. Rad. Isot.* **27**, 431
[17] M Ismail, *Phys. Rev.* **C41**, 87 (1990)
[18] K G Porges, *Phys. Rev.* **101**, 225 (1956)
[19] N T Porile and D L Morrison, *Phys. Rev.* **116**, 1193 (1959)
[20] E A Bryant, D F Cochran and J D Knight, *Phys. Rev.* **130**, 1512 (1963
[21] M Hille, P Hille, M Uhl and W Weisz, *Nucl. Phys.* **A198**, 625 (1970)
[22] H P Graf and H Muenzel, *J. Inorg. Nucl. Chem.* **36**, 3647 (1974)
[23] I A Rizvi, M A Ansari, R P Gautam, R K Y Singh and A K Chaubey, J
3135 (1987)
[24] H D Bhardwaj, A K Gautam and R Prasad, *Pramana – J. Phys.* **31**, 10
[25] N L Singh, B J Patel, D R S Somayajulu and S N Chintalapudi, *Prama*
(1994)
[26] M Blann and H Vonach, *Phys. Rev.* **C28**, 1475 (1983)
[27] M Blann, CODE ALICE/85/300, UCID-20169 (1984)
[28] E Browne and R R Firestone, *Table of radioactive isotopes* (John Wiley
1986)
[29] W D Meyers and W J Swiatecki, *Nucl. Phys.* **81**, 1 (1966)
[30] M Blann and A Mignerey, *Nucl. Phys.* **A186**, 245 (1972)
[31] E Gadioli, E Gadioli-Erba and G Tagliaferri, *Riv. Nuovo Cimento* **6**, 1
[32] E Gadioli, E Gadioli-Erba and J J Hogan, *Phys. Rev.* **C16**, 1404 (1977)

PRAMANA
—journal of
physics

© Printed in India

Projection method for 4d non-orthogonal hype

S JAYANTHI, R K MANDAL* and S LELE
Centre of Advanced Study in Metallurgy, *School of Materials Science and Tec
Technology, Banaras Hindu University, Varanasi 221 005, India

MS received 1 May 1997

Abstract. The Wigner–Seitz cell of a lattice in n-dimensional space displa
group of such a lattice. The vertices of the cell when projected onto pseudo s
outer shape of acceptance domain or motif. This general procedure leads to a
motif identical to those discussed in literature for primitive orthogonal hyper

Example of 4d non-orthogonal hyperlattices corresponding to 12-fold
considered. It will be shown that the first Wigner–Seitz cell degenerates into m
2d pseudo space and can serve as a natural partition of the motif. Following a p
consequence of projection of first 4d Brillouin zone will also be discussed.

Keywords. Wigner–Seitz cell; quasicrystals; non-orthogonal hyperlattices.

PACS Nos 61.44; 61.50

1. Introduction

The two dimensional (d) quasicrystals enjoy a special status in quasic
owing to the presence of periodicity parallel to symmetry axis and q
plane perpendicular to it. Pentagonal [1,2], octagonal [3], dec
dodecagonal [6,7] phases belong to this category. The rank of these
to index their diffraction patterns, one requires five integrally indepe
There have been attempts to understand the nature of two
quasiperiodicity by adopting the method of section/projection [8–1

starting point for the acceptance domain. Such a procedure when space readily gives information about the first Brillouin zone shap and its consequence on physical space projection are explored. We taking the example of a 4d dodecagonal lattice.

2. Determination of Wigner–Seitz cell and acceptance domain

As indicated earlier, the 4 basis vectors in 2d physical and pseudo to recover the metric tensor for dodecagonal lattice given by Brown

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\sin\theta & \cos\theta \\ -\cos\theta & -\sin\theta \end{bmatrix} \begin{bmatrix} x^{\parallel} \\ y^{\parallel} \end{bmatrix},$$

$$\begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix} \begin{bmatrix} x^{\perp} \\ y^{\perp} \end{bmatrix},$$

where V_is and W_is (for $i = 1$ to 4) are respectively the physical a vectors; $\theta = 30°$; $X^{\parallel}$ and $Y^{\parallel}$ are Cartesian bases in physical space v those for pseudo space. The Cartesian bases are complementary other. Thus the direct sum of the above two equations gives base they are

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -\sin\theta & \cos\theta & -\sin\theta & -\cos\theta \\ -\cos\theta & -\sin\theta & \cos\theta & -\sin\theta \end{bmatrix} \begin{bmatrix} x^{\parallel} \\ y^{\parallel} \\ x^{\perp} \\ y^{\perp} \end{bmatrix}$$

where A_is ($i = 1$ to 4) are 4d lattice basis vector. The metric tenso

Table 1. Coordinates of plane normals and equation of hyperplanes.

m_1	m_2	m_3	m_4	Equation of hyperplane
± 1	0	0	0	$\pm(x^{\parallel} + x^{\perp}) = 2$
0	± 1	0	0	$\pm(y^{\parallel} + y^{\perp}) = 2$
0	0	± 1	0	$\pm(-x^{\parallel} + \sqrt{3}y^{\parallel} - x^{\perp} - \sqrt{3}y^{\perp})$
0	0	0	± 1	$\pm(\sqrt{3}x^{\parallel} - y^{\parallel} + \sqrt{3}x^{\perp} - y^{\perp}) =$
± 1	0	± 1	0	$\pm(-x^{\parallel} + \sqrt{3}y^{\parallel} + x^{\perp} - \sqrt{3}y^{\perp})$
0	± 1	0	± 1	$\pm(-\sqrt{3}x^{\parallel} + y^{\parallel} + \sqrt{3}x^{\perp} + y^{\perp})$

Table 2. Coordinate of the vertices of WS cell in terms of 4d cartesian axes in direct and reciprocal space.

Direct*				Reciprocal*			
b_1	b_2	b_2	b_1	b_1	b_1	$\bar{b}_2$	$\bar{b}_2$
b_3	b_4	b_5	$\bar{b}_4$	$\bar{b}_4$	b_3	$\bar{b}_4$	$\bar{b}_5$
$\bar{b}_4$	b_5	b_4	b_3	$\bar{b}_2$	$\bar{b}_2$	b_1	b_1
$\bar{b}_6$	b_6	b_6	$\bar{b}_6$	$\bar{b}_5$	b_3	b_3	$\bar{b}_4$
b_2	b_2	b_1	b_1	b_2	$\bar{b}_2$	$\bar{b}_1$	b_1
b_1	b_1	b_2	b_2	b_4	$\bar{b}_5$	b_4	b_3
b_2	$\bar{b}_1$	b_1	b_2	b_1	$\bar{b}_1$	$\bar{b}_2$	b_2
b_1	$\bar{b}_1$	b_2	$\bar{b}_2$	b_3	b_4	$\bar{b}_5$	b_4
b_1	$\bar{b}_2$	b_2	$\bar{b}_1$	b_5	b_4	$\bar{b}_3$	b_4
b_2	$\bar{b}_1$	b_1	$\bar{b}_2$	$\bar{b}_1$	b_2	b_2	$\bar{b}_1$
b_5	b_4	b_3	$\bar{b}_4$	b_4	b_5	b_4	$\bar{b}_3$
b_2	$\bar{b}_2$	b_1	$\bar{b}_1$	b_6	b_6	b_6	$\bar{b}_6$
b_5	$\bar{b}_1$	b_3	b_1	b_2	b_1	$\bar{b}_1$	$\bar{b}_2$
$\bar{b}_3$	$\bar{b}_1$	b_5	b_1	b_2	b_1	b_1	$\bar{b}_2$
b_1	b_5	b_1	$\bar{b}_3$	b_4	b_3	b_4	$\bar{b}_5$
$\bar{b}_1$	b_3	b_1	$\bar{b}_5$	b_6	$\bar{b}_6$	b_6	$\bar{b}_6$
b_6	$\bar{b}_6$	$\bar{b}_6$	$\bar{b}_6$	b_1	b_2	b_2	$\bar{b}_1$
b_4	$\bar{b}_3$	b_4	$\bar{b}_5$	b_4	$\bar{b}_3$	b_4	b_5

$$b_1 = \frac{3-\sqrt{3}}{6};\ b_2 = \frac{3+\sqrt{3}}{6};\ b_3 = \frac{3-2\sqrt{3}}{6};\ b_4 = \frac{\sqrt{3}}{6};$$

$$b_5 = \frac{3+2\sqrt{3}}{6};\ b_6 = \frac{2\sqrt{3}}{6}.$$

3. 4d reciprocal basis and Brillouin zone

The reciprocal basis vectors of the 4d direct counterparts give
obtained by utilizing the criterion $A_i.A_j^* = \delta_{ij}$ (for $i, j = 1$ to 4). T
delta is 1 for $i = j$ otherwise zero. The reciprocal space basis ve
expressed by

$$\begin{bmatrix} A_1^* \\ A_2^* \\ A_3^* \\ A_4^* \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta & \sin\theta & \cos\theta \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x^{\parallel} \\ y^{\parallel} \\ x^{\perp} \\ y^{\perp} \end{bmatrix}$$

with the identity $[A][A^*]^T = [I]$.

Following a procedure parallel to that explained in the precedi
of the first Voronoi cell have been determined. The 36 vertices thu
table 2. The first Brillouin zone hypersurfaces thus obtained
dodecagons in physical and pseudo reciprocal spaces. We kno
reciprocal lattice points (twice of the Brillouin zone vertex vect
physical space. This means that 4d-Brillouin zone boundaries are
all may not be observable under experimental conditions. This fac
while considering the Fermi surface and Brillouin zone interacti
comprehending the behaviour of electronic states near the Fermi

4. Conclusions

We have presented a method of selecting acceptance doma
hyperlattices. This has been accomplished by first constructing the
then projecting the vertices onto pseudo space. A parallel procedu
construct the 4d-Brillouin zone boundary. The consequence of pro
physical and pseudo space led us to the conclusion that 4d-Wigne
into three dodecagons. This further meant that the 4d-Brillouin zo
in view of diffraction. We believe that such a conclusion
understanding the nature of electronic states near the Fermi surf

[5] L Bendersky, *Phys. Rev. Lett.* **55**, 1461 (1985)
[6] T Ishimasa, H U Nissen and Y Fukano, *Phys. Rev. Lett.* **55**, 511 (1985)
[7] Y Ishimasa and H U Nissen, *Philos. Mag.* **A58**, 835 (1988)
[8] T Janssen, *Acta Cryst.* **A42**, 261 (1986)
[9] J Q You and T B Hu, *Phys. Status Solidi* **B147**, 471 (1988)
[10] U D Kulkarni, S Banerjee and H D Kulkarni, *Acta Met. et Mater.* **41**, 1283
[11] H Brown, R Bülow, J Neubüser, H Wondratschek and H Zassenhans, *Crystal of four dimensional space* (Wiley, New York, 1978)
[12] A Gomez, J L Aragon and F Davila, *J. Phys.* **A24**, 493 (1991)
[13] R K Mandal, *Phys. Rev.* **B50**, 13225 (1994)

PRAMANA
—journal of
physics

© Printed in India

Four constant potential function of amorphous evaluation of its thermodynamic and acoustic

R VENKATESH*, R V GOPALA RAO and SUMITA BANDYOPA
Department of Chemistry, Jadavpur University, Calcutta 700 032, India
*Banaras Hindu University, Varanasi 221 005, India

MS received 9 May 1996; revised 29 November 1996

Abstract. Litov and Anderson after various considerations suggested a f function for a-Se as well as a-As_2S_3. Hence we also used a four constant pote sole purpose of applying this potential function to obtain several acoustic, the properties. We calculated several acoustic properties of a-Se like second c (SOECs), their pressure derivatives, the longitudinal and transverse Grüne different methods, phonon frequencies, absorption band position through th Delaunay's equation, and the thermodynamic properties like heat capacity, Grüneisen constant, the pressure derivative of the bulk modulus (dK_T/d derivative of C_1 which is related to Anderson–Grüneisen parameter, Grüneisen constant namely γ_g'which is related to second Grüneisen const phonon frequencies, potential energy function through the use of fitted para elastic constants. Finally we calculated K_T at the reduced density of ρ/ρ_0 from the potential function with the fitted parameters. In all the above case are found to be in good agreement with experiment wherever available. I important to point out that we eliminated 'C' a constant in the potenti equilibrium condition as was done by Litov *et al* in a-Se and Gerlich *et al* in all amorphous substances are isotropic as mentioned by several author calculate several other properties for a-Se and a-As_2S_3 and present them at

Keywords. Second order elastic constant (SOECs); pressure derivative longitudinal, transverse and thermal Grüneisen constant; absorption band p bulk modulus; pressure derivative of Grüneisen constant; characteristics o potential energy characteristics; third order elastic constants (TOECs)

on a-Se. Extensive description of the properties of amorphous
Elliott [6]. It consists of hexagonal rings and chains. The fact th
that atoms in the chains are strongly bound whereas the interactic
of Van der Walls type [7]. According to Martin *et al* [11] the ap
pressure promotes transfer of electrons from the intra-chain bonc
states between chains take place and can bring about a change
distribution between inter-chain and intra-chain bonds and her
change with pressure [11]. Further, the studies of McCann and Ca
a phase transformation at about 130 kbar when the non-bondin
anti-bonding states at different chains and this causes the weak
bond. Hence the simple Lennard–Jones type of potential funct
various properties theoretically. Therefore one has to add an additi
This leads one to add another term to get theoretically several pr
et al [13] suggested the necessity to add some more terms as
functions fails to give experimental results correctly and several o
extra term. Like others [3, 5], we also added an exponential term -
Jones type of potential function. Here R_0 is the interchain dist
conditions. The interchain distance according to various workers
[14, 15] and is nearer to Van der Walls diameters of 4.0 Å [5].

Litov and Anderson [7] did considerable experimental work c
constant potential function. They experimentally studied the ela
the very well-known condition of equilibrium [16] to eliminate on
potential function.

In the present theoretical investigation we used exactly the
function and eliminated one of the constants as was done by Litov
method of using the equilibrium condition is very well-known s
[16] to eliminate an unknown constant. Litov *et al* [5] and the
same thing. In the present investigations we theoretically obtainec
SOECs by two different methods and calculated the pressure der
as well as the second pressure derivative of bulk modulus and her
the pressure derivative of the Grüneisen constant, the absorpt
various other thermodynamic and acoustic properties. Thus our m
potential function but to see how well it works for predicting the a

Further A, B, C, n, m and λ are potential energy constants. For conve
$m = 12$ and $n = 6$ as was done by many others [16, 7]. If we have t
constants also we require more data. However as was done by Litov and
also assume 12 and 6 as stated above for m and n respectively [
eliminated C [3, 7] by using the equilibrium condition. This of cour
known procedure [16, 17]. This is especially true for all amorphous
from equation (1) we eliminate C as was done by Litov and Anderso
equilibrium condition that $(\mathrm{d}\phi/\mathrm{d}R)_{R=R_0} = 0$ and obtain

$$C = \frac{(mA - nB)\mathrm{e}^{\lambda}}{\lambda}.$$

This was also done by Litov and Anderson in all their investigations
known the bulk modulus K_T is given by

$$K_T = V\left(\frac{\mathrm{d}^2\phi}{\mathrm{d}V^2}\right) = \left(\frac{\mathrm{d}^2\phi}{\mathrm{d}V^2}\right)V$$

which can be rewritten for amorphous substances as [16]

$$K_T = \frac{N}{9V}\left[R^2\,\frac{\mathrm{d}^2\phi}{\mathrm{d}R^2}\right]_{R=R_0}.$$

We define a dimensionless quantity,

$$\rho/\rho_0 = V_0/V = 1/v = (R_0/R)^3,$$

where we assumed along with others [3, 7, 17] that for amorphous subst

From equations (1), (3) and (4) we get

$$K_T = \frac{N}{9V_0}\left[156A\left(\frac{1}{v}\right)^4 - 42B\left(\frac{1}{v}\right)^2 - (12A - 6B)\lambda v^{2/3}\mathrm{e}^{\lambda}(1 -\right.$$

Here V_0 is the volume of a-Se under a pressure of one atmosphere.
important to discuss the nature of the roots of equation (1). If we have
the right hand side and with $m = 12$ and $n = 6$ we will have 12 roots bu
of an exponential term it becomes transcendental type of equation

while $-C\,e^{-(\lambda R/R_0)}$ is positive and hence repulsive. There is absc having two attractive terms and one repulsive term namely $-C$ discussion can be found where two attractive forces exists as fluoride [16]. Thus $m = 12$ and $n = 6$ are just chosen values and give the first two terms as attractive while third term which is an e $-C\,e^{-\lambda R/R_0}$ as repulsive. This is just a fit with three K_T values an λ are not unique. In what follows we give very large number of p these parameters. However the potential function with these par same type of potential function as obtained by Litov and Anders they also form another set of parameters. It may further be point simple alkali metals one has to use two repulsive terms as was Reed and Gubins [19] give an extensive discussion on var functions. In fact the third term $-C\,e^{(\lambda R/R_0)}$ which is repu mechanically acceptable than any bireciprocal function [17].

2.2 *Evaluation of certain acoustic properties of a-Se with the u*

Bhatia and Singh (BS) defined two quantities β and δ as

$$\beta = \frac{\rho_w R^2}{2M}\left[\frac{1}{R}\frac{\mathrm{d}\phi}{\mathrm{d}R}\right]_{R=R_0}$$

and

$$\delta = \frac{\rho_w R_0^3}{2M}\left[\frac{\mathrm{d}}{\mathrm{d}R}\left(\frac{1}{R}\frac{\mathrm{d}\phi}{\mathrm{d}R}\right)\right]_{R=R_0}.$$

Here M is the mass of the atom, ρ_w is the weight density and R two nearest chains while R_0 is its equilibrium value. Since w $(\mathrm{d}\phi/\mathrm{d}R)_{R=R_0} = 0$, and hence $\beta = 0$ as can be seen from equatio

$$\delta = \frac{\rho_w R_0^2}{2M}\left(\frac{\mathrm{d}^2\phi}{\mathrm{d}R^2}\right)_{R=R_0}$$

vide equation (7).

From equation (7) it can be easily shown that

$$\frac{\mathrm{d}\delta}{\mathrm{d}P} = \frac{2}{3K_T} - \frac{\rho_w R_0^3}{6K_T M}\left(\frac{\mathrm{d}^3\phi}{\mathrm{d}R^3}\right)_{R=R_0}.$$

In the derivation of equation (13) it is assumed as was done by Gerlich amorphous substances, the volume $V \propto R^3$ unlike in crystalline Se. From can easily get

$$\left(\frac{\mathrm{d}^3\phi}{\mathrm{d}R^3}\right) = \frac{m(m+1)(m+2)}{R^3}\left(\frac{R}{R_0}\right)^{-m} + \frac{Bn(n+1)(n+2)}{R^3}\left(\frac{R}{R_0}\right)^{n} + \left(\frac{mA - Bn}{R_0^3}\right)\lambda^2 \mathrm{e}^{[1-(R/R_0)]\lambda}.$$

Using equations (11) to (14) we can calculate $(\mathrm{d}C_{11}/\mathrm{d}P)$ and $(\mathrm{d}C_{44}/\mathrm{d}P)$ [21, 22] using Schofields equation [23] that

$$\left(\frac{\mathrm{d}C_{11}}{\mathrm{d}P}\right)_T = \frac{C_{11}}{K_T} + 1.8(C_1^T - 1),$$

$$\left(\frac{\mathrm{d}C_{44}}{\mathrm{d}P}\right)_T = \frac{C_{44}}{K_T} + 0.6(C_1^T - 1).$$

It was shown that [24]

$$C_1^{\mathrm{L,T}} = 2\gamma_{\mathrm{g}}^{\mathrm{L,T}} + \tfrac{1}{3}.$$

This equation is a generalization by the well known Debye's equation,

$$C_1 = 2\gamma_{\mathrm{g}}^{\mathrm{th}} + \tfrac{1}{3}.$$

Here $\gamma_{\mathrm{g}}^{\mathrm{L}}$ and $\gamma_{\mathrm{g}}^{\mathrm{T}}$ are the longitudinal and transverse Grüneisen constants f possible to calculate the same from Bhatia–Singh (BS) parameters [22]

For isotropic substances like a-Se they are shown to be (22)

$$\frac{\mathrm{d}C}{\mathrm{d}P} = \left[\frac{6}{5} + \frac{2}{5}\frac{L_1 S}{}\right]$$

$$C_1 = \frac{dK_T}{dP} \equiv \frac{1}{3} - \frac{NR_0^3}{27K_T V}\left(\frac{d^3\phi}{dR^3}\right)_{R=R_0}.$$

Using the potential parameters we get $C_1 = 6.95$ while the exper
This can be considered as good agreement. From the value of
$C_1' \equiv (dC_1/dP)$ from (14) and (20). Thus the final expression fo
algebra can be obtained as

$$C_1' = \frac{1}{K_T}\left[\frac{NR_0^3}{27K_T V_0}\left\{C_1\phi'''(R) + \frac{R_0}{3}\phi''''(R)\right\}\right].$$

The fourth derivative can be obtained from equation (14).
rearranged to give

$$C_1' = \frac{1}{K_T}\left[\frac{32760NA}{81K_T V_0} - \frac{3024NB}{81K_T V_0} - \frac{C\lambda^4 N}{81K_T V_0}e^{-\lambda} - C_1^2\right.$$

Here N is Avagadros number and V_0 is molar volume and is equ
The value obtained for $C_1' = 0.118\,(\text{kbar})^{-1}$ (and hence $\gamma_g' = 0$
This is a small value and unless one is very sure of the accu
calculations can be of great help. This is related to Anderson Gr

Another simple method of calculating γ_g' is from the well kn

$$\gamma_g = -\frac{a}{6}\left[\frac{\phi'''(R)}{\phi''(R)}\right]_{R=R_0}.$$

Here as pointed earlier, a is the interatomic distance between two
chain

$$\gamma_g^{th} = -\frac{1}{6}a[\phi'''(R)(\phi''(R))^{-1}]_{R=R_0},$$

$$\frac{d\gamma_g^{th}}{dP} = -\frac{1}{6}a\left[\frac{\phi''''(R)}{\phi''(R)} - \left(\frac{\phi'''(R)}{\phi''(R)}\right)^2\right]_{R=R_0}\left(\frac{dR}{dP}\right)_{R=R_0}$$

We assume as done by Gerlich *et al* and others [2, 5, 7] $V \propto R^3$

$$\gamma_g' = -\frac{1}{6}a\left[-\frac{\phi''''(R)R}{\phi''(R)3K_T} + \frac{R}{3K_T}\left(\frac{\phi'''(R)}{\phi''(R)}\right)^2\right]_{R=R_0},$$

$$\gamma_g' = \frac{a}{K_T}\left[\frac{\phi''''(R)}{\phi''(R)}\frac{R_0}{18} - \frac{2R_0a^2}{36a^2} - \left(\frac{\phi'''(R)}{\phi''(R)}\right)^2\right].$$

Using the values of A, B, λ with $C = ((mA - nB)/\lambda)\,e^{+\lambda}$ we get $\gamma_g' =$
We used $a = 2.32$ Å (5), while the value obtained from equation (22
These values can be considered as good since the values are very sm
experimental error will cause this discrepancy.

2.4 *Evaluation of the band absorption position and bulk modulus*

From the calculated values of SOECs obtained from the potenti
equations (9) and (10) it is possible to calculate the absorption band po
Smith De-Launay's equation [27] which is given by

$$\frac{8A_1[A_1 + 8C_{11} - 16C_{44}]}{[3A_1 - 8C_{11} + 16C_{11}^2]} = 1,$$

where

$$A_1 = \omega^2 M/2a.$$

Here $a = 2.32$ Å and is the nearest neighbour distance [28, 5] betwee
in the same chain and ω is the absorption frequency. We evaluate th
position since $(1/\lambda) = \omega/2\pi C_0$. Here C_0 is the velocity of light. We c
modulus from equation (31) and (32). Thus differentiating equations
respect to pressure and rearranging we get

$$\left\{8\left(\frac{dC_{11}}{dP}\right)[3A_1 - 8C_{11} + 16C_{12}] + 16\left(\frac{dC_{12}}{dP}\right)(8C_{11} - 16\right.$$

$$\times (A_1 - 24C_{11} + 16C_{12})^{-1} = \frac{A_1}{K_T}[2\gamma_g^{el} + \tfrac{1}{3}].$$

2.5 *Computation of thermal Grüneisen parameter*

From the potential function it is possible to calculate the therma
and is given by the well known equation [29]

$$\gamma_g^{th} = -\frac{1}{6}\, a \left[\frac{\phi'''(R)}{\phi''(R)}\right]_{R=R_0}.$$

Here R_0 is the interchain distance which is the van der Waals
4.0 Å [7]. The value obtained is given in table 3.

Another method of calculating γ_g^{th} is as follows. We have fro

$$\frac{d^2\phi(R)}{dR^2} = \phi''(R) = \frac{Am(m+1)R_0^m}{R^{m+2}} - \frac{Bn(n+1)R_0^n}{R^{n+2}} -$$

We may point out that $[\phi''(R)]_{R=R_0}$ can be treated as the force
value the heat capacity of a-Se was evaluated which gives a valu
with experiment. Hence we have

$$\omega^2 = \left[\frac{\phi''(R)}{M}\right]_{R=R_0}.$$

Here M is the atomic mass and ω is the frequency.

$$2\ln\omega = \ln[-\phi''(R)] - \ln M$$

$$2\left(\frac{\partial \ln\omega}{\partial \ln V}\right) = \frac{\partial \ln[\phi''(R)]}{\partial \ln V}.$$

By definition

$$\left(\frac{\partial \ln\omega}{\partial \ln V}\right) = -\gamma_g^{th}.$$

Hence from equation (37) we get,

$$-2\gamma_g^{th} = \frac{\partial \ln[\phi''(R)]}{\partial \ln V},$$

$$\frac{\partial \ln \phi''(R)}{\partial \ln R} \qquad \frac{\partial \ln R}{\partial \ln R} \qquad \frac{\partial \ln R}{\partial \ln R}$$

justified. Thus it can be considered that the various values of γ_g are in goo with experiment obtained by different methods.

2.6 *Calculation of phonon frequency from the potential function*

We can calculate the phonon frequencies using Bhatia–Singh's metho equations for longitudinal and transverse phonon frequencies of Bhatia an given by [20]

$$\rho_w \omega_L^2 = \frac{2n_c}{a^2}(\beta_0 + \delta I_2) + \frac{K_e K_{TF}[G(Kr_s)]^2}{K^2 + K_{TF}^2 \tilde{g}(R)},$$

$$\rho_w \omega_T^2 = \frac{2n_c}{a^2}[\beta I_0 + 0.5\delta(I_0 - I_2)].$$

Here n_c is the number of nearest neighbours, K_e is the electronic bulk modu Thomas–Fermi screening length, $G(Kr_s)$ is the shape factor and the rest of have their usual significances [22]. The electronic bulk modulus for a sem substance like a-Se K_e is zero. Hence we do not want to further delineate r second term in equation (43). The terms I_0 and I_2 are defined as

$$I_0 = 1 - \frac{\sin x}{x},$$

$$I_2 = \frac{1}{3} - \left[\frac{1}{x} - \frac{2}{x^3}\right]\sin x - \frac{2\cos x}{x^2}$$

and

$$x = ka.$$

The calculated phonon frequencies of $w_L(K)$ and $w_T(K)$ are shown in fi

2.7 *Calculation of third order elastic constants (TOECs)*

There are no experimental values of TOECs and hence the th
important. We use the equations of Rao and Venkatesh [30], T
[32]. These are given by

$$\frac{C_{111}}{C_{11}} = -\left[5 + \frac{C_{11}}{K_T}(C_1 - 1)\right],$$

$$\frac{dC_{11}}{dP} = C'_{11} = -\frac{1}{3K_T}(C_{111} + 2C_{112}),$$

$$C'_{44} = -\frac{1}{3K_T}(C_{144} + 2C_{166}),$$

$$C'_{12} = -\frac{1}{3K_T}(C_{112} + C_{123}).$$

In order to calculate C_{166} we use Shiren's [31] non-linear p
$\langle 110 \rangle$ directions and

$$K = 3 + \frac{C_{111}}{C_{11}},$$

$$K = 3 + \frac{1}{\alpha_1}[1/2C_{111} + 1.5C_{112} + 6C_{166}],$$

where

$$\alpha_1 = C_{11} + C_{12} + 2C_{44}$$

and K is Shiren's constant.

Here for a-Se since no direction can be fixed as it is isotropic
equations (52) to (54) are applicable in the calculation of K. Th
(54) we calculate C_{456} through the well known equation [9]

$$C_{166} = C_{144} + 2C_{456}.$$

In this connection it may be emphasized that for isotropic s
three independent TOECs. However, for completeness we ha
TOECs. In order to check the veracity of the calculated TOEC w

Substituting the values of A, B and λ along with assumed values of done by Litov and Anderson [7] whose procedure was followed totally i value of $C = (10.5726/\lambda)\,e^{+\lambda} \times 10^{-12}$ erg we get,

$$\phi''(R_0) = \frac{24.60}{R_0^2} \times 10^{-12}.$$

As stated earlier $\phi''(R)$ can be treated as a force constant and hence

$$\omega^2 = \frac{\phi''(R)}{M}.$$

We get $\omega = 1.2376 \times 10^{13}\,\text{sec}^{-1}$ with $R_0 = 4.0\,\text{Å}$ since we are calcul temperature. We used Einstein's equation which is given by [34]

$$C_v = \frac{3Ry^2e^y}{(e^y - 1)^2}.$$

Here

$$y = \frac{\hbar\omega}{K_B T},$$

where the symbols have their usual significances.

The calculated value of C_v with the parameters necessary for eva $5.95\,\text{cal}/^\circ\text{K}$ gm-atom while the experimental value is $5.83\,\text{cal}/^\circ\text{K}$ g agreement being excellent.

3. Results and discussion

A large number of properties of a-Se isotropic substance have been e method of Litov and Anderson [7]. Even very complex amorphous sub treated as isotropic and it is extremely difficult to give the large nu references. Since the potential function is a inverse 12th and 6th powe exponential term added to it, the set of roots are not unique, neither unphysical in having either two repulsive or two attractive terms [35,

Table 2. Several calculated properties of a-Se with the potenti

Property	Calc. values	Expt. values	Eq
dC_{11}/dP	15.2	14.0	
dC_{44}/dP	3.3	2.7	
$\gamma_g^L(S)^*$	2.9	3.5	1
$\gamma_g^T(S)^*$	2.7	3.0	1
γ_g^L (B–S)*	4.2	3.5	
γ_g^T (B–S)*	3.3	3.0	
A_1	579.4 (GPA)	—	
C_v	5.95 (cal/°K gm atom)	583 (cal/°K gm atom)	
$1/\lambda$	240 cm^{-1}	254 cm^{-1}	
K_T	8.7 (GPA)	9.1 GPA	3

*S = From Schofield's equations, B–S= Bhatia–Singh's param

Table 3. Thermal Gruneisen's constant γ_g^{th}.

γ_g^{th} value Calc.	γ_g^{th} Expt.	Equations used	Ref. for expt. value
1.38	0.96	24	[26]
1.23		42	[26]

Table 4. Calculated values of C_1' and γ_g'.

C_1'	γ_g'	Equations used
0.170	0.085	30
0.118	0.059	22

$C_1 = 2\gamma_g'$ (vide equation 17).

We present in table 1, C_{11} and C_{44} and they are in very good a Brassington *et al* [9].

Table 5. Characteristics of phonon frequencies (vide figure 1).

	Position		Peak height in
	From P.F.	From E.C.	From P.F.
ω_L	1.78 Å^{-1}	1.80 Å^{-1}	2.01
ω_T	2.50 Å^{-1}	2.46 Å^{-1}	1.59

P.F. = Potential function; E.C. = Elastic constants.

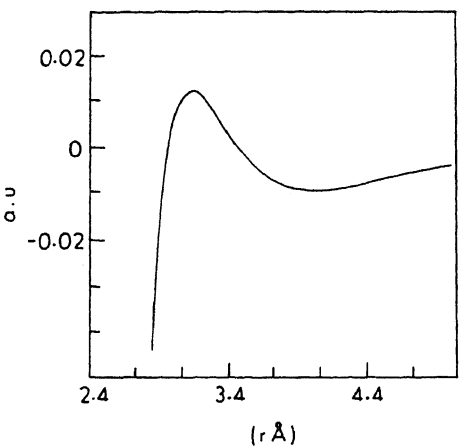


Figure 2. Potential function of a-Se.

Table 6. Potential energy characteristics.

Potential energy	Calculated	Litov and Anders
Maximum position	3.2 Å	3.29 Å
Value	0.013 a.u.	–
Minimum position	4.0 Å	4.0 Å
Value	−0.0097 a.u.	–

Table 7. Calculated TOECs.

TOECs	Values in GPA
C_{111}	−228.0
C_{112}	−64.8
C_{123}	−105.2
C_{144}	+6.9
C_{166}	−40.3
C_{456}	−23.6

Another important property which can be obtained from the heat capacity with the aid of the potential parameters. The agree value and the experiment is excellent. We would like to point o (5) we calculate K_T at $\rho/\rho_0 = 1.1 = V_0/V \equiv 1/\nu$. Substituting we obtain a value of 139 kbar for k_T while the experimental val shows the veracity of the parameters. Several other acoustic ar calculated which will be presented later on along with various o have calculated extensively both acoustic and thermodyna agreement in most of the cases is very good and this gives parameters which produce exactly the same type of potential Litov and Anderson [7].

Acknowledgements

The authors are thankful to the Indian National Science Academ and Technology and Council of Scientific and Industrial R financial assistance.

References

[1] H Hoshino, R W Schmutzler and F Hansel, *Third International* (Institute of Physics, Bristol, 1977) p. 404

[13] R M Martin, T A Fjeldy and W Richter, *Solid State Commun.* **18**, 86
[14] K Vedam, D L Miller and R Roy, *J. Appl. Phys.* **37**, 3432 (1966)
[15] A Van Hippel, *J. Chem. Phys.* **16**, 372 (1948)
[16] Moelwyn Hughes, *Physical Chemistry* (Pergamon Press, Oxford and N
322, p. 556, pp 336–339
[17] Max Born and Kun Huang, *Dynamical theory of crystal lattices* (O
1968) pp 22–25
[18] F Seitz, *The modern theory of solids* (McGraw Hill Book Company,
1940) p. 381
[19] T M Reed and K Gubbins, *Applied Statistical Mechanics* (McGraw Hi
York and London, 1973) ch. 4
[20] A B Bhatia and R N Singh, *Phys. Rev.* **B31C**, 4751 (1985)
[21] R V Gopala Rao and R Venkatesh, *J. Non-Cryst. Solids* **155**, 245 (19
[22] R V Gopala Rao and R Venkatesh, *Pramana – J. Phys.* **43**, 201 (199
[23] P Schofield, *Proc. Phys. Soc.* **88**, 149 (1966)
[24] S E Babb (Jr), *J. Chem. Phys.* **38**, 27–43 (1963)
[25] F Y Hansen and T S Knudsen, *J. Chem. Phys.* **62**, 1556 (1975)
[26] R V Gopala Rao and R Venkatesh, *J. Phys. Chem. Solids* (1996) (to
[27] J De-Launay, *Solid State Physics* **2**, 296 (1956); **7**, 301 (1958)
[28] Y Waseda, K Yokoyama and K Suzuki, *Phys. Condens. Matter* **18**, 2
[29] P P M Meincke, *Can. J. Phys.* **40**, 283 (1962)
[30] R V Gopala Rao and R Venkatesh, *Phys. Rev.* **B39**, 9467 (1989)
[31] R N Thurston, *Proc. of IEEE* **53**, 132 (1965)
[32] M Pomerantz, *Phys. Rev.* **A139**, 501 (1965)
[33] F Birch, *Phys. Rev.* **71**, 809 (1947)
[34] A O E Animalu, *Intermediate quantum theory of crystalline solids* (Pr
Ltd., New Delhi, 1977) p. 93
[35] F Seitz, *The modern theory of solids* (McGraw Hill Book Compar
London, 1940) p. 381
[36] G Perrin and Delaunay-Coutris, *J. Phys. Chem. Solids* **49**, 1397 (198
[37] M P Brassington, A J Miller and G A Saunders, *Philos. Mag.* **B43**,

PRAMANA
—journal of
physics

© Printed in India

Vol
Sep
pp.

High pressure study of phase transitions in α-FePO

CHITRA MURLI, SURINDER M SHARMA, S K KULSHRESHTHA* and
High Pressure Physics Division, *Chemistry Division, Bhabha Atomic Research Cen
Mumbai 400 085, India

MS received 24 April 1997; revised 14 July 1997

Abstract. High pressure behaviour of $FePO_4$ in berlinite form has been investigated using vibrational Raman spectroscopy and energy dispersive x-ray diffraction. Co these techniques along with studies on pressure quenched samples reveal structural this material from its room pressure trigonal phase to a disordered and a crystallin 3 ± 0.5 GPa. The latter is the Cmcm phase which is the equilibrium structure at hi These high pressure phases do not revert back to its initial structure after release Irreversibility of these transformations indicates that FeO_4 tetrahedra do not regai coordination. These high pressure transitions can be rationalized in terms of the th energy diagram for such systems.

Keywords. α-$FePO_4$; phase transition; x-ray diffraction; Raman scattering.

PACS Nos 62.50; 61.50; 64.70

1. Introduction

High pressure behaviour of materials having structures based on tetrahe work are of interest because of their relevance to the geophysical prob structures display a large number of phase transitions depending on compressibility of the constituent tetrahedra and the nature of inter-tetrahed [1]. Also the tetrahedral framework does not favour close packing and t effect of pressure in these structures provides an insight into relative im

connection α-$FePO_4$ which is isostructural to quartz family of interest.

The difference in the cation–O interactions manifests itself parameters, such as cation–O–cation angle, at ambient conditio SiO_2 and α-$AlPO_4$ indicate that phase transformation as well as a a decrease in the non-bonded O—O distances due to the red (T = Si, Al, Fe, P, etc.) (see ref. [4] and references therein.) The that if the initial T–O–T angle is smaller then these phase transf at lower pressures. These angles are, T–O(1)–$T = 144°, 143.1°$ $T = 144°, 142°$, and 137.5° for α-SiO_2, α-$AlPO_4$ and α-FePO suggests that α-$FePO_4$ will undergo pressure induced phase pressures.

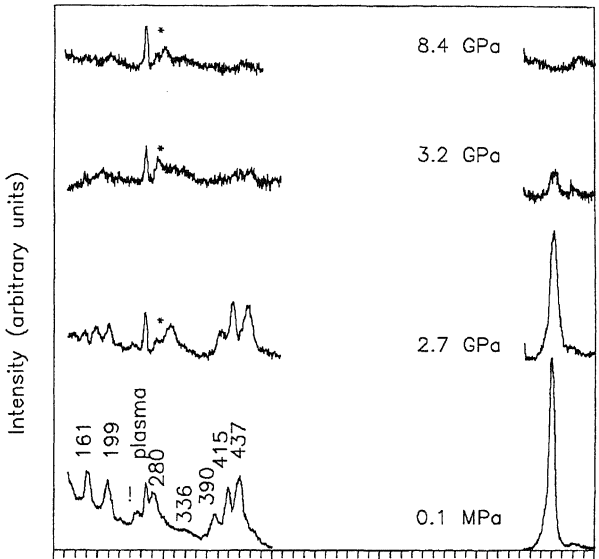
2. Experimental details

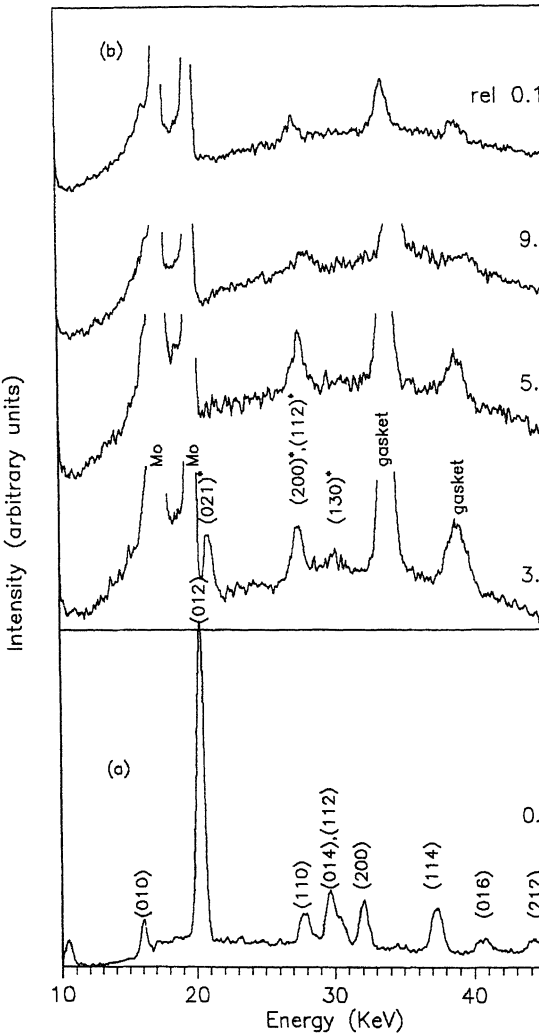
The berlinite form of $FePO_4$ belongs to the space group $P3_121$ per unit cell. The cell constants of the sample used [11] $c = 11.245$ Å which compare well with the results of Ng *et a* $c = 11.255$ Å). The condensed polycrystallites were used for I were powdered for x-ray diffraction studies. High pressure exp using fluorescence free Raman quality diamonds (whose second signal to noise ratio of > 2) with culet sizes of $\sim 500\,\mu$ set in a cell. The sample was loaded along with a ruby chip of size 10–20 in a steel gasket, preindented to a thickness of $\sim 100\,\mu$. Methan ratio of 4 : 1 was used as pressure transmitting medium, and bridging of the sample between the diamond tips. The pressure well-known ruby fluorescence technique.

High pressure Raman scattering experiments were carried out uses a single-stage double-pass 500 mm Chromex scanning mo notch filter from Kaiser optics. This filter permits the observatio than 100 cm^{-1}. We used the 514.5 nm line of a 150 mw Ar-ion power as the excitation source and some of its plasma lines for

Table 1. Comparison of Raman modes of α-$FePO_4$ and α-$AlPO_4$.

α-$FePO_4$		α-$AlPO_4$	
ν_i	$d\nu_i/dP$	ν_i	$d\nu_i/dP$
A_1 modes		A_1 modes	
199	0.1	160	0.0
280	1.5	216	2.6
436	0.6	460	1
1018	0.3	1110	−0.5
E modes		E modes	
161	0.65	116	0.8
336	—	372	−0.1
390	0.45	416	0.4
415	0.4	436	0.9




(b)
rel 0.1
9.
5.
3.
Mo
Mo
(021)*
(012)
(200)*,(112)*
(130)*
gasket
gasket
(a)
0.
(010)
(110)
(014),(112)
(200)
(114)
(016)
Intensity (arbitrary units)
10
20
30
40
Energy (KeV)

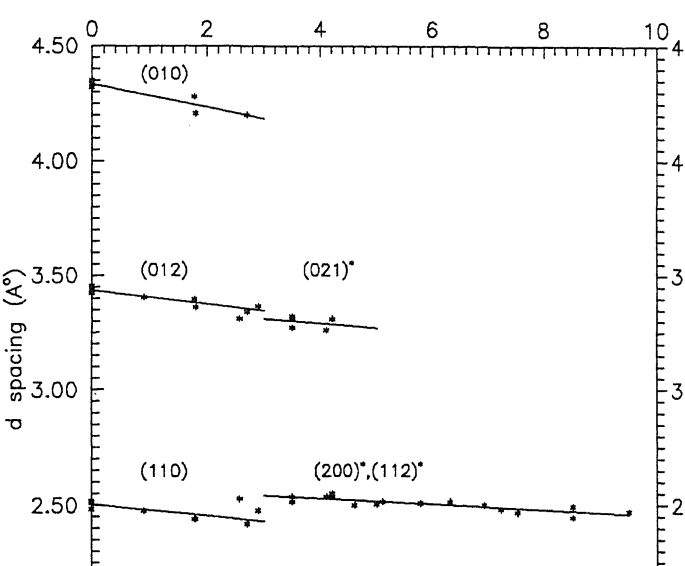


Figure 3. The *d* values of first three Bragg lines of berlinite phase as pressure. Also shown are lines (*) of the Cmcm phase-data comp different experiments. The spread in *d* values at room pressure gives t measurements.

Figure 2 shows the x-ray diffraction pattern of α-$FePO_4$ at different pr diffraction pattern of the recovered sample shows that it does not revert berlinite phase, in contrast to memory glass behaviour observed in α-$AlPO_4$. of the Bragg peaks decrease with pressure up to 3 ± 0.5 GPa (see figur pressure the strong (012) peak disappears and a weaker peak emerges close a new broad line appears near the original (110) peak. These changes indicat

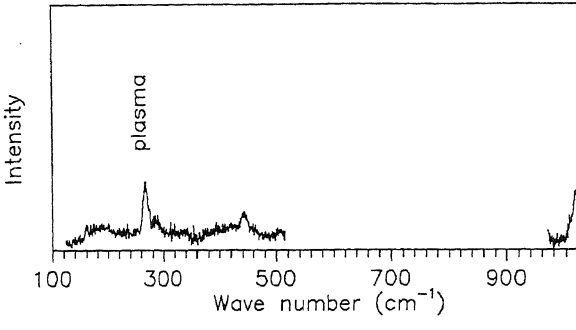


Figure 4. Raman spectrum of pressure quenched sample.

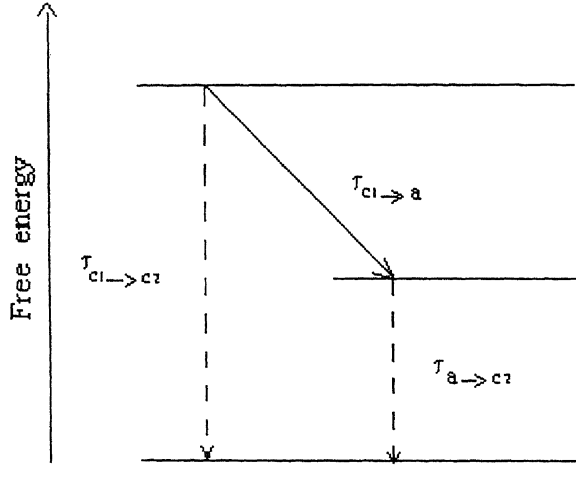


Figure 5. A schematic three level diagram for $c \rightarrow a$ tran

the samples were brownish black in colour and its Raman spec
background, which we believe is due to amorphization. The x-ra

that the energy barriers for $c_1 \rightarrow a$ or $a \rightarrow c_2$ are of similar mag
rule out the direct transformation of $c_1 \rightarrow c_2$.

Our experimental data show that c/a ratio of α-$FePO_4$ increa
around 2 GPa, before it undergoes the phase transition. This fe
what is observed in this class of materials [1]. As anticipated
chemistry, the pressure at which the compounds of this family exh
correlated with the ambient values of T–O–T angle, referred to

In most of these compounds, it has been shown, through IR al
as well as molecular dynamics calculations, that crystal t
transformations are closely related to coordination increase as
In the present case the presence of Cmcm phase in the pressure rel
that Fe atoms which is six coordinated to oxygens in the Cmcm p
initial four coordination to neighbouring oxygens.

References

[1] R M Hazen and L W Finger, *Phase Trans.* **1**, 1 (1979); *Am. Sci.* **72**
110 (1985)
W G Moffat, G W Pearsall and J Wulff, *The structure and prop*
Eastern Limited, New Delhi, 1980) vol. 1, Ch. 5

[2] S K Sikka, Surinder M Sharma and R Chidambaram, *High pressu*
edited by S C Schmidt, J W Shaner, G A Samara and M Ross (A

[3] R J Hemley, C T Prewitt, K J Kingma, *Silica, reviews in minerology*
Prewitt and G V Gibbs (Minerological Society of America, Wash
p. 241

[4] Surinder M Sharma and S K Sikka, *Progress in materials scie*
references therein

[5] K J Kingma, R J Hemley, H K Mao and D R Veblen, *Phys. Rev.*

[6] M S Somayazulu, Nandini Garg, Surinder M Sharma and S K Sikk
(1994)

[7] H Sankaran, Surinder M Sharma, S K Sikka and R Chidambaram,
(1990)

[8] M B Krugger and R Jeanloz, *Science* **249**, 647 (1990)

[9] P Cordier, J C Doukhan and J Peyronneau, *Phys. Chem. Miner.* **2**

[10] P Gillet, J Badro, B Varrel and P F Macmillan, *Phys. Rev.* **B51**, 1

[11] M M Gadgil and S K Kulshreshtha, *J. Solid. State. Chem.* **111**, 3

PRAMANA
—journal of
physics

© Printed in India

Thermoelectric power of high-T_c superconductors extended Hubbard model

S KESHRI and P K BARHAI
Department of Applied Physics, Birla Institute of Technology, Mesra, Ranchi 835

MS received 2 May 1997

Abstract. The thermoelectric power (TEP) for a one dimensional lattice has b using the extended Hubbard model in the limit $U \neq \infty$, where U is the on-site Coul A new expression for TEP, derived in this study, has been found to successfully experimental results of the insulating samples of different hole- and electr temperature superconducting systems.

Keywords. Hubbard model; thermoelectric power.

PACS No. 74.72

1. Introduction

Hubbard model [1] basically deals with the theory of correlation effect metals having narrow d-bands. This model has been successfully used metal-insulator (M-I) transition behaviour of transition metals. Hubbard m been extensively used for the analysis of transport properties of many conc organic compounds, e.g., *N*-methylphenazinium-tetracyanoquinodime TCNQ) etc. [2, 3] and high-temperature superconducting (HTSC) sampl of which show M-I transitions [7, 8].

In some earlier papers [9a, 10], we have reported the method of measur analysis of TEP of HTSC samples by using Hubbard model and also other e

reasonably well on the basis of Hubbard model but marked dev
concentrations.

(2) For some other hole-doped HTSC systems, e.g., Tl-2212 [1
between theoretical and experimental values are observed
concentrations.

These findings show that Hubbard model as used in these stu
success. One of the shortcomings of this model is the neglect
electron repulsion in the Hamiltonian given by eq. (1), which, a
Kwak and Beni [12] in the study of TEP of conducting salts of org
NMP-TCNQ etc., has significant influence on the TEP of these san
of the influence of the intermolecular electron repulsion, Kwak
modified the Hamiltonian given by eq. (1) to the form

$$H = U\sum_i n_{i\uparrow}n_{i\downarrow} + V\sum_i n_i n_{i+1} - t\sum_{i,\sigma}(C^{\dagger}_{i,\sigma}C_{i+1,\sigma} + \mathrm{H}$$

where V represents the intermolecular electron repulsion. This
Hamiltonian has often been referred to as the extended Hubbard m
that has been used in our present study of TEP of HTSC samples. H
the study of Kwak and Beni [12] is restricted to conducting salts of
therefore our first objective is to use the Hamiltonian of eq. (2)
samples. The logic behind this investigation is that since both NI
samples show M-I transitions [7, 8] we expect the extended Hubba
Beni [12] to be equally applicable in the study of TEP of HTSC s
study differs from Kwak and Beni in the limits of U. In the analy
[12], to reduce the complexity of calculations, Kwak and Beni ha
$U = \infty$ for the study of TEP of a Hubbard chain with arbitrary
restricted their study to bands less than half-filled so that $n_i = 0$ o
use the extended Hubbard model in the limit $U \neq \infty$. The physic
limit is to include also the cases for which n_i may have values oth
knowledge this study represents the first investigation on TEP o
extended Hubbard model in the limit $U \neq \infty$.

Our investigation shows that although the above form of the Ha
(2), is successful in explaining the TEP behaviour of the insulatin

experimental data of TEP of the insulating samples of different hole- and el
HTSC systems.

2. Thermoelectric power

Following the method of Beni [2] and restricting to a one-dimensional latt
Kubo's formulas [15] for the transport coefficient of a many-body system, T
defined [2] as

$$S = \frac{-S^{(2)}/S^{(1)} + \mu/e}{T},$$

where e is the absolute value of electronic charge, μ, the chemical potenti

$$S^{(1)} = \frac{1}{2}\beta e^2 \int_0^\infty \langle v\, v(\tau) + v(\tau) v \rangle \mathrm{d}\tau$$

and

$$S^{(2)} = \frac{1}{2}\beta e \int_0^\infty \langle Q v(\tau) + v(\tau) Q \rangle \mathrm{d}\tau$$

are the correlation functions with

$$v(\tau) = \mathrm{e}^{-i\tau H} v \mathrm{e}^{i\tau H}.$$

In eqs (4) and (5), $\beta = (kT)^{-1}$, k is the Boltzmann constant, T is
temperature, v and Q are the velocity and the energy-flux operators respecti
represents thermal averages. The operators v and Q are given by

$$v = \operatorname*{Lt}_{q\to 0} \frac{1}{\hbar q} \sum_j [n_j, H] \mathrm{e}^{\mathrm{i}q(j\alpha)},$$

and

$$Q = \operatorname*{Lt}_{q\to 0} \frac{1}{\hbar q} \sum_j [h_j, H] \mathrm{e}^{\mathrm{i}q(j\alpha)},$$

where the t^2-terms have been dropped, as usually done in the narrov
Using the same approximation, the full Hamiltonian H in $S^{(1)}$ and S
the unperturbed Hamiltonian

$$H_0 = U\sum_i n_{i\uparrow}n_{i\downarrow} + V\sum_i n_i n_{i+1}.$$

Using eqs (4)–(6), the expression for $S^{(2)}/S^{(1)}$ has been found to

$$\frac{eS^{(2)}}{S^{(1)}} = \frac{\int_0^\infty d\tau\, \mathrm{Tr}[e^{-\beta(H_0-\mu\sum_i n_i)}(Qe^{-iH_0\tau}ve^{iH_0\tau} + e^{-iH_0\tau}$$
$$\int_0^\infty d\tau\, \mathrm{Tr}[e^{-\beta(H_0-\mu\sum_i n_i)}(ve^{-iH_0\tau}ve^{iH_0\tau} + e^{-iH_0\tau}$$

The traces of the above equation may be obtained in a straight fo

$$e^{-iH_0\tau}C_{j,\sigma}e^{iH_0\tau} = e^{i\tau\{Un_{j,-\sigma}+V(n_{j-1,\sigma}+n_{j+1,\sigma})\}}C_{j,\sigma}.$$

Considering only four adjacent sites $|i-1, i, i+1, i+2\rangle$ for the ca
and after some lengthy algebra we obtain from eq. (12)

$$\frac{eS^{(2)}}{S^{(1)}} = U\{x^2e^{-\beta U-2\beta V} + x^3(e^{-2\beta U-5\beta V} + e^{-\beta U-4\beta V}) + x$$
$$\{1 + x(1+e^{-\beta U-\beta V}) + x^2(e^{-\beta U-\beta V} + e^{-\beta U-3\beta V}$$
$$+ x^3(e^{-2\beta U-4\beta V} + e^{-\beta U-4\beta V} + x^4e^{-2\beta U-6\beta V})\},$$

where $x = e^{\beta\mu}$. Using (14) in (3) we obtain,

$$S = -\frac{k}{e}[\beta U\{x^2e^{-\beta U-2\beta V} + x^3(e^{-2\beta U-5\beta V} + e^{-\beta U-4\beta V})$$
$$\{1 + x(1+e^{-\beta U-\beta V}) + x^2(e^{-\beta U-\beta V} + e^{-\beta U-3\beta V})$$
$$+ x^3(e^{-2\beta U-4\beta V} + e^{-\beta U-4\beta V} + x^4e^{-2\beta U-6\beta V})\} - \ln$$

This is the expression for TEP of extended Hubbard model in the
be noted that if one uses $V = 0$ in eq. (15), the final expression r
the narrow-band Hubbard chain at arbitrary electron density
(eq. (12), [21].

where $\rho = N_e/N$ is the electron density and N_e is the number of electrons. F
and (17) we obtain

$$x = \frac{-(1-\rho)e^{-\beta V} + \{(1-\rho)^2 e^{-2\beta V} + (\rho/2)(2-\rho)e^{-\beta U - 2\beta V}\}^{1}}{(2-\rho)e^{-\beta U - 2\beta V}}$$

This equation, together with eq. (15), completes the determination of S.

We now use the above results for the analysis of TEP of HTSC sample
samples $U \cong 5\,\text{eV}$ [9b, 14, 16]. This shows that U is very large for th
compared to the value of kT at room temperature ($kT = 0.026\,\text{eV}$ at $T = 3$
use $U \gg kT$ in eqs (15) and (18), they reduce to

$$S = \frac{k}{e} \ln x$$

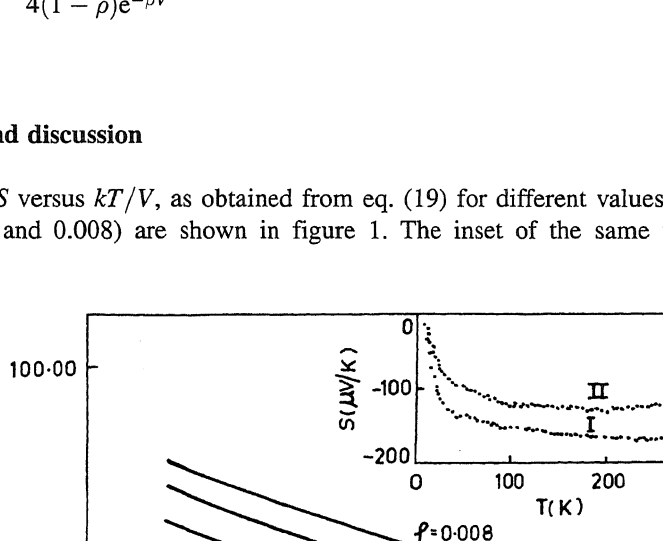
and

$$x = \frac{\rho}{4(1-\rho)e^{-\beta V}}$$

respectively.

3. Results and discussion

The plots of S versus kT/V, as obtained from eq. (19) for different values
0.004, 0.006 and 0.008) are shown in figure 1. The inset of the same



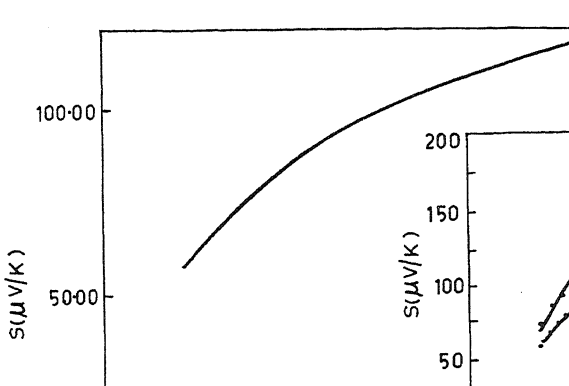


Figure 2. The variation of S with $kT/|V|$ for $\rho = 0.962$, ca
The curves A of inset shows the experimental
$Tl_2Ba_2Ca_{0.5}Y_{0.5}Cu_2O_{8+y}$ sample (reproduced from figure 4

the experimental curves of S versus T of figure 1 of Xu *et al* samples $Nd_{1.99}Ce_{0.01}CuO_4$ and $Nd_{1.975}Ce_{0.025}CuO_4$ of the electro $Nd_{2-x}Ce_xCuO_4$. It is seen that S versus kT/V curves of f experimental curves of inset of figure 1. This shows that eq. expression for S in the limit $U \neq \infty$, is successful to produce the the insulating samples of $Nd_{2-x}Ce_xCuO_4$ system. However, we ha other data of temperature variation of TEP of electron-doped samp density, to test its general validity.

We have also used eq. (19) for the explanation of experime the insulating samples of our hole-doped HTSC systems, Bi_2Sr and $Tl_2Ba_2Ca_{1-x}Y_xCu_2O_{8+y}$ [10]. The experimental data (figu

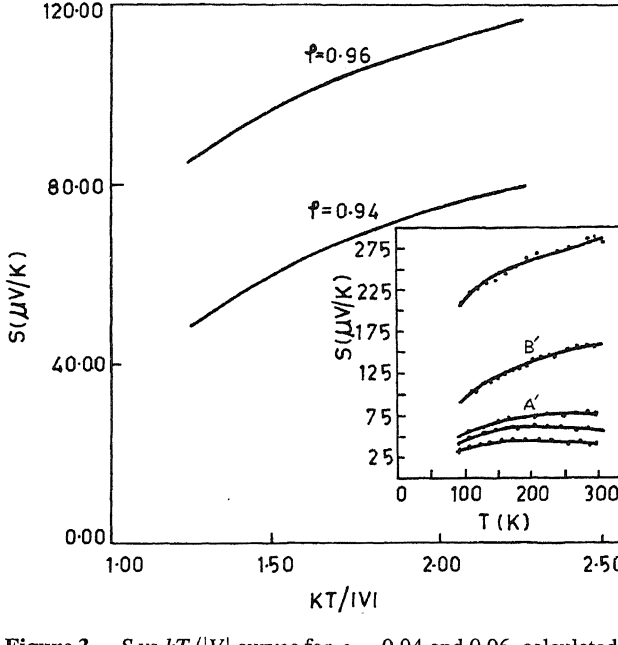


Figure 3. S vs $kT/|V|$ curves for $\rho = 0.94$ and 0.96, calculated
curves A′ and B′ of the inset represent the experimental c
$Bi_2Sr_2Ca_{0.4}Y_{0.6}Cu_2O_{8+y}$ and $Bi_2Sr_2Ca_{0.3}Y_{0.7}Cu_2O_{8+y}$ sample
duced from figure 4(b) of Mandal *et al* [9a]).

The experimental results of TEP (reproduced from figure 4(
insulating samples $Bi_2Sr_2Ca_{0.4}Y_{0.6}Cu_2O_{8+y}$ (with $p = 0.06$, i
$Bi_2Sr_2Ca_{0.3}Y_{0.7}Cu_2O_{8+y}$ with ($p = 0.04$, i.e. $\rho = 0.96$), are shown in
3 by the curves A′ and B′ respectively. The plots of S versus $kT/|V$
0.96, as obtained from eq. (19) with negative V, are shown in figure
the plots of figure 3 with the plots of the inset of figure 3 show
temperature region, eq. (19) with negative V is also successful to re

References

[1] J Hubbard, *Proc. R. Soc. (London)* **A276**, 238 (1963)
[2] G Beni, *Phys. Rev.* **B10**, 2186 (1974)
[3] G Beni and F Coll III Cornelius, *Phys. Rev.* **B11**, 573 (1975)
[4] P W Anderson, *Science* **235**, 1196 (1997)
[5] G Baskaran, Z Zou and PW Anderson, *Solid State Commun.* **63**, 973 (1987)
[6] A N Das, B Ghosh and P Choudhury, *Physica* **C158**, 311 (1989)
[7] A J Epstein, S Elemad, A F Garito and A J Heeger, *Phys. Rev.* **B5**, 952 (1972)
[8] B Ellman, H M Jager, D P Katz, T F Rosenbum, A S Cooper and G P Espinosa, *Phys. Rev.* **B39**, 9012 (1989)
[9a] J B Mandal, S Keshri, P Mandal, A Poddar, A N Das and B Ghosh, *Phys. Rev.* **B46**, 11840 (1992)
[9b] P Mandal, A Poddar, P Choudhury, J B Mandal, S Keshri, A N Das and B Ghosh, *Indian J. Pure Appl. Phys.* **30**, 531 (1992)
[10] S Keshri, J B Mandal, P Mandal, A Poddar, A N Das and B Ghosh, *Phys. Rev.* **B47**, 9048 (1993)
[11] J R Cooper, B Alavi, L W Zhou, W P Beyermann and G Grüner, *Phys. Rev.* **B35**, 8794 (1987)
[12] J F Kwak and G Beni, *Phys. Rev.* **B13**, 652 (1976)
[13] X-Q Xu, S J Hazen, W Jiang, J L Peng, Z Y Li and R L Greene, *Phys. Rev.* **B45**, 7356 (1992)
[14] Mandal, A Poddar, B Ghosh and P Choudhury, *Phys. Rev.* **B43**, 13102 (1991)
[15] Kubo, *J. Phys. Soc. Jpn.* **12**, 1203 (1957)
[16] S-W Cheong, S E Brown, Z Fisk, R S Kwok, J D Tompson, E Zirngiebl, G Grüner, D E Peterson, G L Wells, R B Schwarz and J R Cooper, *Phys. Rev.* **B36**, 3913 (1987)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 3,
September 1997
pp. 301–308

Radially dependent antishielding factor $\gamma(r)$ of the Sc^{3+} ion in the solid state

B B PANIGRAHI, N C MOHAPATRA[1] and S HAFIZUDDIN[2]
Department of Physics, Khallikote (Autonomous) College, Berhampur 760 001, India
[1]Department of Physics, Berhampur Univeristy, Berhampur 760 001, India
[2]Department of Physics, B.J.B. College, Bhubaneswar 751 001, India

MS received 13 November 1996; revised 16 July 1997

Abstract. We report the results of our calculation of $\gamma(r)$, the radially dependent antishielding factor of Sc^{3+} ion in the crytalline environment. Watson sphere model is used to approximately represent the crystal potential. Differential equations resulting from non-orthogonal Hartree–Fock perturbation theory are solved to obtain perturbations to core electron states. Contribution to $\gamma(r)$ from electron self-consistency effect has been evaluated by using the latter wave-functions in the many-body expression of Schmidt *et al.*

Keywords. Radially dependent antishielding factor; crystalline potential; electron-electron self consistency effect.

PACS Nos 32.10; 31.70

1. Introduction

It is well known [1–3] that when source charges, such as conduction electrons in metals, are inside the core region of the atoms, antishielding effect of the core states becomes radially dependent. If the perturbation of the core states is by the nuclear quadrupole moment Q, the antishielding factor is referred [2] to as $\gamma(r)$. The latter plays an important role in the calculation of electric field gradient (EFG) in metals. In most of the early calculations [4–6] of EFG in metals, $\gamma(r)$ has been set equal to zero primarily because the latter quantities were not available at the time of these calculations. In a few cases [7, 8] where $\gamma(r)$ values were available, these were not used in EFG calculations perhaps for two reasons: First these were obtained for a free ion instead of an ion in solid state. Secondly, the calculated results of $\gamma(r)$ did not include the so called electron self-consistency contribution which, in general, is not negligible.

With the introduction of Watson sphere model [9], it has now been possible to approximately include the crystalline effect on $\gamma(r)$. As to the consistency contribution, it has been shown by Ray *et al* [10] as well as Beri *et al* [11] that the total consistency contribution $\gamma_\infty^{(1)}$ can be evaluated from their many-body expression using the perturbation of the core electron states obtained in differential equation procedures [2, 12]. Following their [10, 11] suggestions, Schmidt *et al* [13] have calculated $\gamma_\infty^{(1)}$ for a large number of positive and negative ions both in free and crystalline states. A further

extension has been subsequently made by Tripathi and Mohapatra [14] for evaluating $\gamma^{(1)}(r)$, the radially dependent consistency contribution, from the expression of $\gamma_\infty^{(1)}$ originally due to Ray *et al* [10].

We have followed closely the procedure of Tripathi and Mohapatra [14] for evaluating $\gamma(r)$ in Sc^{3+} ion in solid state. The reasons for taking up Sc are two-fold: First $\gamma(r)$ for this ion is not available either in free state or in the solid state. Secondly, we wish to use these results in the calculation of EFG in Sc metal to study the change it may produce in the EFG obtained recently [15, 16] with an approximate $\gamma(r)$.

We wish to remark here that there are a number of procedures in which EFG can be evaluated without using $\gamma(r)$ explicitly. These are: (a) linearized augmented plane wave method [17, 18], (b) linear muffin-tin-orbital method [19, 20], (c) augmented spherical wave method [21], (d) linear combination of atomic orbital method [22] and the procedure due to Meyer *et al* [23]. In the latter procedure [23], various schemes are developed to reconstruct the true wave functions from the pseudo wave functions in crystals and EFG is evaluated with good accuracy using the true wave functions.

The paper is organized as follows: In § 2, we briefly describe the theory. Results and discussion are presented in § 3. Section 4 summarizes the conclusions.

2. Theory

If consistency contribution is neglected, the resulting antishielding factor, referred to as the zero-order contribution and denoted by $\gamma^{(0)}(r)$ is given [14] by

$$\gamma^{(0)}(r) = \frac{1}{Q}\left(\int_0^r Q_i^{(0)}(r')\mathrm{d}r' + r^5 \int_r^\infty Q_i^{(0)}(r')(r')^{-5}\mathrm{d}r'\right), \tag{1}$$

where

$$Q_i^{(0)}(r) = Qr^2 \sum_n C_{l_0 l}^{(2)} U_0^l(nl, r) V_1^l(nl \to l_0, r). \tag{2}$$

In eq. (2) U_0^l is the radial part of the unperturbed core state. V_1^l is the radial part of the perturbation in the core state and is obtained in the differential equation procedure of non-orthogonal Hartree–Fock perturbation theory (NHFPT) of Dalgarno [24]. Other symbols have their usual meanings. Following the procedure of Tripathi and Mohapatra [14], the radially dependent consistency contribution can similarly be expressed as

$$\gamma^{(1)}(r) = \frac{1}{Q}\left(\int_0^r Q_i^{(1)}(r')\mathrm{d}r' + r^5 \int_r^\infty Q_i^{(1)}(r')(r')^{-5}\mathrm{d}r'\right), \tag{3}$$

where $Q_i^{(1)}(r)$ is interpreted as the induced quadrupole moment density in the core electrons due to the consistency effect. In obtaining consistency contribution [10], one considers the simultaneous excitation of a pair of electrons. Correspondingly, one may have contributions from radial, angular and mixed modes of excitations. In the radial mode, both the core electrons are excited into states with unchanged l values. In the angular mode both the core electrons are excited into states with $l \pm 2$ values. In the mixed mode while one is excited to a state with same l, the other goes to $l \pm 2$ state. For convenience, we will refer to the sum of angular and the mixed mode as the 'angular mode' contribution.

It has been noted by Schmidt *et al* [13] that the contribution to $\gamma_\infty^{(1)}$ from the 'angular mode' is quite negligible compared to the radial mode. In the present calculation we have neglected the former.

For p-core ions, the expressions for $Q_i^{(1)}(r)$ have been worked out by Tripathi and Mohapatra [14]. These expressions, in addition to depending on U_0', V_1' functions also involve $\overline{V}_1'$, the radial part of the perturbed core states in EFG-perturbation. For details the reader may refer to the work of Tripathi and Mohapatra [14].

3. Results and discussion

Hartree–Fock (HF) wave functions of ions in free state as well as in crystalline state, the latter being evaluated in Watson sphere model [9] are available from the work of Paschalis and Weiss [25]. The core wave functions of Sc^{3+} ion in crystalline state corresponding to Pauling [26] ionic radius, $r_0 = 0.81$ Å, are used in the present work. These are taken from the work of Paschalis and Weiss [25]. Out of the five core states, $1s-3p$ of Sc^{3+}, the s-core states contribute to $\gamma^{(0)}(r)$ through only the angular mode of excitation, the two p-core states contribute both in radial and angular modes. Out of these $\gamma_{\text{ang}}^{(0)}(r)$, the net angular mode contribution to $\gamma^{(0)}(r)$ is generally small and negligible compared to the more important radial mode contribution, $\gamma_{\text{rad}}^{(0)}(r)$, but the former takes more computation time than the latter. For this reason, $\gamma_{\text{ang}}^{(0)}(r)$ has been evaluated using a modified Thomas–Fermi model for the charge distribution in exactly the same way as it was done in the previous work [12].

The radial mode contribution $\gamma_{\text{rad}}^{(0)}(r)$, on the other hand, has been evaluated exactly by using the appropriate HF wave function $U_0'(2p,r)$ and $U_0'(3p,r)$ of $2p$ and $3p$ core electrons. The differential equations in the perturbations $V_1(np \rightarrow p,r)$ and $\overline{V}_1(np \rightarrow p,r)$ are solved using a non-iterative procedure due to Duff and Das [27]. These functions are then orthogonalized to the occupied core states in order to avoid violation of exclusion principle. In the numerical solution of the differential equations in V_1 and $\overline{V}_1$, a logarithmic scale with 209 mesh points has been chosen. The r-value of the nth mesh point in this scale is obtainable from the relation, $r_n = Z^{-1}\exp[-3.0 + (n-1)h]$, where Z is the atomic number of Sc and $h = 0.045$. The largest value of r in this choice of mesh points is $r_{209} = 27.536$ atomic units (au).

In figures 1 and 2, we have plotted the orthogonalized solutions V_1' for the $2p$ and $3p$ core states of Sc^{3+} ion in solid state. These have the same number of nodes as of a $4p$ atomic core state in agreement with the expectation from perturbation of $2p$ and $3p$ states in the radial modes. In figure 3 the total $\gamma^{(0)}(r)$, that is, the sum of $\gamma_{\text{rad}}^{(0)}(r)$ and $\gamma_{\text{ang}}^{(0)}(r)$ is plotted as a function of r. The nature of its variation with r leading to a nearly constant value at large r is as expected.

The consistency contribution $\gamma^{(1)}(r)$, has been evaluated using the expressions derived in the work of Tripathi and Mohapatra [14]. In doing so the contributions from the radial mode have been calculated exactly. On the other hand, the contribution from the 'angular mode' which is generally smaller than the radial mode, but takes more computation time has been neglected in the present work. The error due to this neglect is negligible. It may be noted here that unlike Na^+ ion [14], which has only $2p$ core electrons, Sc^{3+} ions has both $2p$ and $3p$ core electrons. Therefore, in the latter ion, contribution to $\gamma^{(1)}(r)$

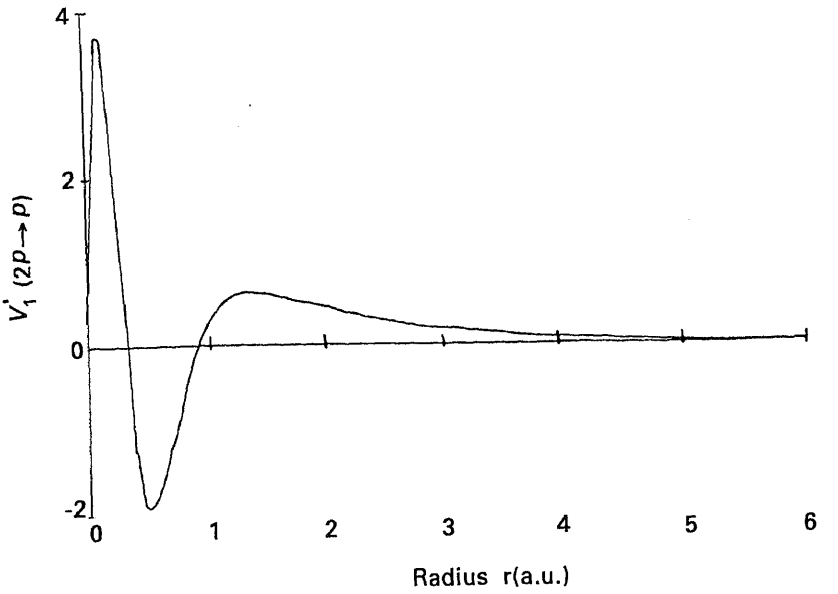


Figure 1. Perturbed function $V_1'(2p \rightarrow p)$ of Sc^{3+} ion in solid state.

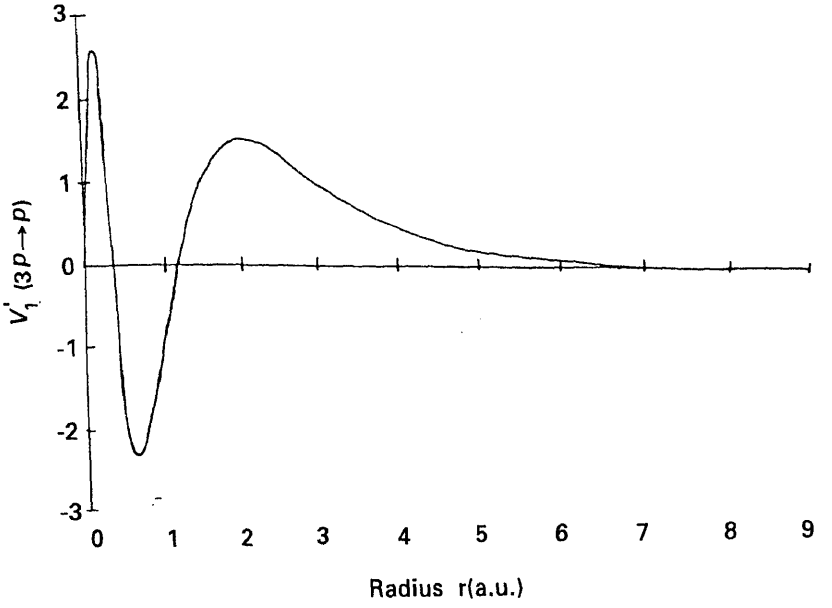


Figure 2. Perturbed function $V_1'(3p \rightarrow p)$ of Sc^{3+} ion in solid state.

arises both from intrashell and intershell pair excitations of p-core electrons. In figure 4 $\gamma^{(1)}(r)$ is plotted. The variation with r is similar to those observed in case of other ions [14].

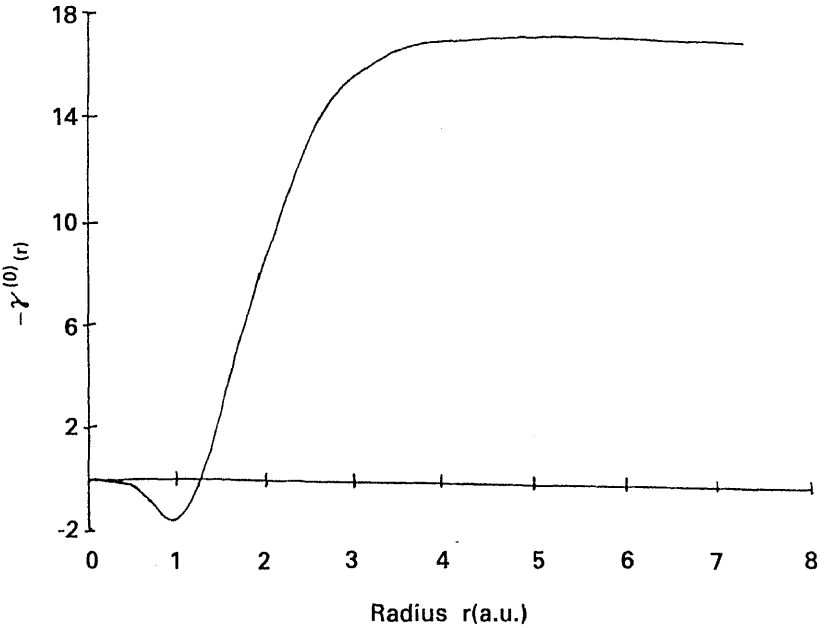


Figure 3. The radially dependent antishielding factor $\gamma^{(0)}(r)$ of Sc^{3+} ion in solid state.

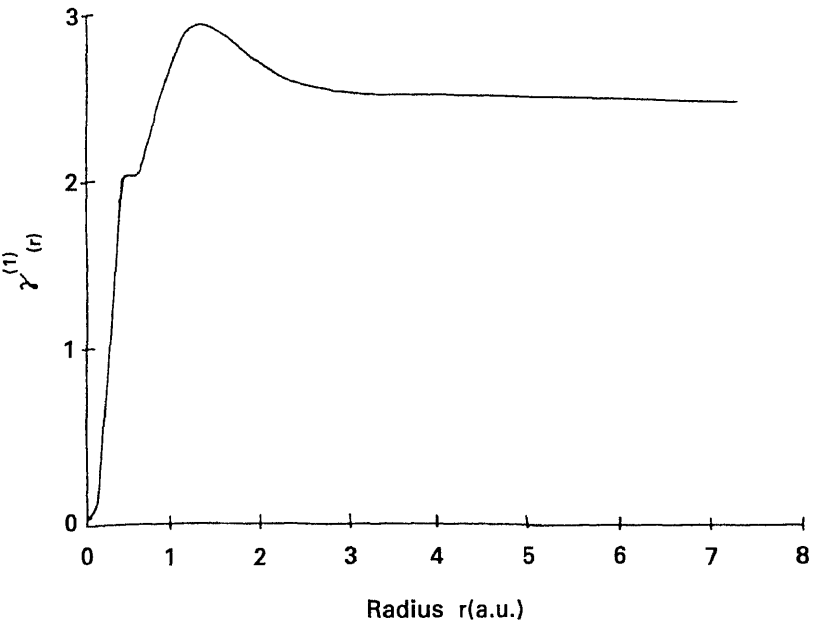


Figure 4. Self-consistency contribution $\gamma^{(1)}(r)$ to radially dependent antishielding factor of Sc^{3+} ion in solid state.

Table 1. The radially dependent antishielding factors $\gamma^{(0)}_{rad}(r)$, $\gamma^{(0)}_{ang}(r)$, $\gamma^{(0)}(r)$, $\gamma^{(1)}(r)$ and $\gamma_{tot}(r)$ of Sc^{3+} ion in solid state. r is expressed in atomic units and n denotes the index of the mesh points.

$n(r)$	r	$\gamma^{(0)}_{rad}(r)$	$\gamma^{(0)}_{ang}(r)$	$\gamma^{(0)}(r)$	$\gamma^{(1)}(r)$	$\gamma_{tot}(r)$
1	0.0024	0.0000	0.0001	0.0001	−0.0000	0.0001
69	0.0505	0.0011	0.0165	0.0176	−0.0501	−0.0325
79	0.0793	0.0067	0.0312	0.0379	−0.2061	−0.1681
85	0.1038	0.0171	0.0456	0.0627	−0.4244	−0.3617
93	0.1488	0.0461	0.0745	0.1206	−0.9288	−0.8082
100	0.2040	0.0725	0.1132	0.1857	−1.5059	−1.3202
105	0.2555	0.0596	0.1516	0.2112	−1.8553	−1.6441
120	0.5018	−0.0997	0.3455	0.2458	−2.0350	−1.7892
124	0.6008	0.1156	0.4232	0.5388	−2.1663	−1.6275
128	0.7193	0.5176	0.5128	1.0304	−2.4176	−1.3872
133	0.9007	0.8775	0.6376	1.5151	−2.7595	−1.2444
136	1.0309	0.6266	0.7177	1.3443	−2.8976	−1.5533
138	1.1280	0.1123	0.7728	0.8851	−2.9437	−2.0586
140	1.2343	−0.7414	0.8287	0.0873	−2.9520	−2.8647
141	1.2911	−1.3018	0.8570	−0.4448	−2.9430	−3.3878
142	1.3505	−1.9535	0.8854	−1.0681	−2.9271	−3.9952
143	1.4127	−2.6871	0.9138	−1.7733	−2.9046	−4.6779
144	1.4777	−3.5029	0.9423	−2.5606	−2.8777	−5.4383
145	1.5457	−4.3829	0.9707	−3.4122	−2.8468	−6.2590
146	1.6168	−5.3261	0.9990	−4.3271	−2.8141	−7.1412
147	1.6913	−6.3089	1.0271	−5.2818	−2.7801	−8.0619
148	1.7691	−7.3307	1.0550	−6.2757	−2.7468	−9.0225
149	1.8505	−8.3639	1.0826	−7.2813	−2.7143	−9.9956
150	1.9357	−9.4063	1.1099	−8.2964	−2.6842	−10.981
151	2.0248	−10.4300	1.1367	−9.2933	−2.6563	−11.950
153	2.2155	−12.386	1.189	−11.197	−2.6099	−13.807
155	2.4241	−14.131	1.239	−12.892	−2.5759	−15.468
156	2.5357	−14.902	1.263	−13.639	−2.5634	−16.203
158	2.7745	−16.202	1.308	−14.894	−2.5453	−17.439
160	3.0358	−17.181	1.350	−15.831	−2.5347	−18.366
163	3.4746	−18.105	1.405	−16.700	−2.5271	−19.227
172	5.2095	−18.775	1.518	−17.257	−2.5243	−19.781
175	5.9624	−18.783	1.539	−17.244	−2.5241	−19.769
179	7.1383	−18.784	1.556	−17.228	−2.5241	−19.752
180	7.4669	−18.784	1.556	−17.228	−2.5241	−19.752
209	27.536	−18.784	1.556	−17.228	−2.5241	−19.752

In addition to these figures, we have also summarized the data in table 1. Instead of listing the values of $\gamma^{(0)}$ and $\gamma^{(1)}$ at all the 209 mesh points, we have chosen for the sake of brevity, to list the values at thirty-five representative mesh points, so that the broad features of $\gamma^{(0)}$ and $\gamma^{(1)}$ (figures 3 and 4), namely, node/hump in small r region and near constancy in the large r region are represented by this smaller set of selected mesh points. The first column gives the n-value of the mesh point while the second column gives the r value corresponding to this n. Columns 3–7 list the contributions $\gamma^{(0)}_{rad}(r)$, $\gamma^{(0)}_{ang}(r)$, $\gamma^{(0)}(r)$, $\gamma^{(1)}(r)$ and $\gamma_{tot}(r)$, the last one being the sum of $\gamma^{(0)}(r)$ and $\gamma^{(1)}(r)$. From this table, the saturation values of $\gamma^{(0)}$, $\gamma^{(1)}$ and γ_{tot} are noted to be −17.23, −2.52 and −19.75

respectively, which may be regarded as equal to the respective Sternheimer antishielding factor $\gamma_\infty^{(0)}$, $\gamma_\infty^{(1)}$ and γ_∞ caused by source charges at infinity.

As $\gamma^{(0)}(r)$ and $\gamma^{(1)}(r)$ from other works are not available, it is not possible to compare the present results with others. However, the values of $\gamma_\infty^{(0)}$, $\gamma_\infty^{(1)}$, and γ_∞ for Sc^{3+} ion in solid state are available from the work of Schmidt *et al* [13]. Their values are: $\gamma_\infty^{(0)} = -20.46$, $\gamma_\infty^{(1)} = -2.644$ and $\gamma_\infty = -23.104$. Comparing these values with those in the present work, we find the present results are smaller in magnitude than those of Schmidt *et al* [13]. The discrepancy between the present results and those of Schmidt *et al* [13] may be attributed to the different ionic radii at which the unperturbed core wave functions are evaluated in Watson sphere model [9]. While the HF wave functions used in the work of Schmidt *et al* [13] correspond to $r_0 = 0.73$ Å, those in the present work refer to Pauling radius $r_0 = 0.81$ Å. In case of positive ions, as can be seen from the work of Schmidt *et al* [13], a larger r_0 gives values for $\gamma_\infty^{(0)}$ and $\gamma_\infty^{(1)}$ with smaller magnitudes. For negative ions, the trend is opposite. In the present work $r_0 (= 0.81$ Å) being larger than 0.73 Å, the latter value used by Schmidt *et al* [13], the present results of $\gamma_\infty^{(0)}$ and $\gamma_\infty^{(1)}$ are smaller in magnitude than theirs, as expected. We would like to add here that contribution arising from higher order electron correlation effect which is expected to be still smaller than consistency contribution has been neglected in the present work.

Another pertinent point to discuss is the relative magnitudes of $\gamma^{(0)}(r)$ and $\gamma^{(1)}(r)$. Considering the source from which $\gamma^{(1)}(r)$ results, one expects it to have smaller magnitude than $\gamma^{(0)}(r)$. This is borne out by the data in table 1 for r larger than about 1.6 au. But for $r < 1.6$ au, $\gamma^{(1)}(r)$ is found larger in magnitude than $\gamma^{(0)}(r)$. In a bid to understand this we have separately listed $\gamma_{\text{rad}}^{(0)}(r)$ and $\gamma_{\text{ang}}^{(0)}(r)$ which make up $\gamma^{(0)}(r)$. In the small r region, $\gamma_{\text{rad}}^{(0)}(r)$ starts from a small positive value, attains a maximum and then decreases and becomes negative. $\gamma_{\text{ang}}^{(0)}(r)$, on the other hand, is positive and gradually increases. In contrast, $\gamma^{(1)}(r)$ is negative and gradually increases in magnitude. This trend is more or less the same as observed in case of Na^+ ion in the work of Tripathi and Mohapatra [14]. We do not have a simple explanation for this. It may be due to the complicated dependence of $\gamma^{(1)}(r)$ on the functions U_0', V_1' and $\bar{V}_1'$ in the small r region. As to the neglected 'angular contribution' to $\gamma^{(1)}(r)$, we expect it to be much smaller than 1.57, the saturation value of $\gamma_{\text{ang}}^{(0)}(r)$. If one assumes that the 'angular contribution' to $\gamma^{(1)}(r)$ scales down in the same ratio as $\gamma_{\text{rad}}^{(1)}(r)$ does in relation to $\gamma_{\text{rad}}^{(0)}(r)$, the saturation value of $\gamma_{\text{ang}}^{(1)}(r)$, the 'angular contribution' will be less than 0.22.

Before we conclude this section we wish to discuss the sensitiveness of $\gamma(r)$ to the Watson sphere radius. This type of study has been made by Schmidt *et al* [13] in Cl^-. In Sc^{3+}, as we have stated before, the value of γ_∞ changes from -19.75 to -23.104 when the radius r_0 changes from 0.81 Å to 0.73 Å. That is, γ_∞ changes by 17% for a change of 10% in r_0. The change in $\gamma(r)$ for individual r may be slightly different from this. The uncertainty in Pauling [26] ionic radius which is used for r_0 in the present work is expected [13] to be less than 10%. This suggests that the change in the present result due to any uncertainty in Pauling ionic radius will be much less than 17%. It is for this reason, we expect that the EFG in Sc metal when calculated by using the present results of $1 - \gamma(r)$ would not significantly change if the actual ionic radius differs from the Pauling radius (0.81 Å) by less than 10%. However, the exact change can be determined if one repeats the calculation of EFG using $\gamma(r)$ corresponding to the aforesaid two values of r_0. Such calculations are planned for future.

4. Conclusion

Radially-dependent antishielding factor $\gamma(r)$ of Sc^{3+} ion in crystalline state has been evaluated in the differential equation procedure of non-orthogonal Hartree–Fock perturbation theory. Consistency contribution from the dominant radial mode of excitation has also been included. The present results of $\gamma(r)$ may be used in the calculation of EFG in Sc metal for obtaining better accuracy than those in other procedures where $\gamma(r)$ is either neglected or poorly approximated.

References

[1] H M Foley, R M Sternheimer and D Tyoko, *Phys. Rev.* **83**, 734 (1954)
[2] R M Sternheimer, *Phys. Rev.* **146**, 140 (1966)
[3] E H Hygh and T P Das, *Phys. Rev.* **143**, 452 (1966)
[4] N C Mohapatra, C M Singal, T P Das and P Jena, *Phys. Rev. Lett.* **29**, 456 (1972)
[5] P Jena, S D Mohanti and T P Das, *Phys. Rev.* **B7**, 975 (1973)
[6] N C Mohapatra, C M Singal and T P Das, *Phys. Rev. Lett.* **31**, 530 (1973)
[7] S K Padhi and N C Mohapatra, *Phys. Rev.* **B29**, 5199 (1984)
[8] B K Acharya and N C Mohapatra, *Pramana – J. Phys.* **43**, 391 (1994)
[9] R E Watson, *Phys. Rev.* **111**, 1108 (1958)
[10] S N Ray, T Lee, T P Das and R M Sternheimer, *Phys. Rev.* **A9**, 1108 (1974)
[11] A C Beri, S N Ray, T P Das and R M Sternheimer, *Phys. Rev.* **A12**, 1168 (1975)
[12] K K Prasad Rao and N C Mohapatra, *Phys. Rev.* **A24**, 10 (1981)
[13] P C Schmidt, K D Sen, T P Das and Al Weiss, *Phys. Rev.* **B22**, 4167 (1980)
[14] P Tripathi and N C Mohapatra, *J. Phys.* **B23**, 3241 (1990)
[15] B B Panigrahi and N C Mohapatra, *Pramana – J. Phys.* **41**, 443 (1993)
[16] B B Panigrahi and N C Mohapatra, *Pramana – J. Phys.* **43**, 407 (1994)
[17] D D Koelling and G O Arbman, *J. Phys.* **F5**, 2041 (1975)
[18] P Blaha, K Schwarz and P H Dederichs, *Phys. Rev.* **B37**, 2792 (1988)
[19] O K Andersen, *Phys. Rev.* **B12**, 3060 (1975)
[20] M Methfessel and S Frota-Pessöa, *J. Phys.: Condens. Matter* **2**, 149 (1990)
[21] A R Williams, J Kübler and C D Gelatt Jr, *Phys. Rev.* **B19**, 6094 (1979)
[22] H Eschrig, *Optimized LCAO Method and the Electronic Structure of Extended Systems* (Springer, Berlin, 1989)
[23] B Meyer, K Hummler, C Elsässer and M Fähnle, *J. Phys.: Condens. Matter* **7**, 9201 (1995)
[24] A Dalgarno, *Proc. R. Soc. (London)* **A251**, 282 (1959)
[25] E Paschalis and Al Weiss, *Theor. Chim. Acta.* **13**, 381 (1967)
[26] L Pauling, *The Nature of Chemical Bond*, 3rd edn (Cornel University Press, Ithaca, NY, 1960)
[27] K J Duff and T P Das, *Phys. Rev.* **168**, 43 (1968)

PRAMANA
— journal of
physics

© Printed in India

Vol. 49, No. 3,
September 1997
pp. 309–316

Time dependent magnetization of a $Zn_{0.1}Fe_{0.9}Fe_2O_4$ fine magnetic particle system

R V UPADHYAY
Department of Physics, Bhavnagar University, Bhavnagar 364002, India

MS received 24 September 1996; revised 13 March 1997

Abstract. The time dependent magnetization measurements of the fine particles (100 Å) of $Zn_{0.1}Fe_{0.9}Fe_2O_4$ ferrite system at 68 K are reported. The time decay of magnetization is found to be logarithmic. The time dependent co-efficient varies with the applied field and is maximum at coercive field. The value of anisotropy obtained from the variation of time dependent co-efficient with field agrees well with earlier results.

Keywords. Fine particles; anisotropy.

PACS Nos 75.50; 75.60

1. Introduction

A magnetic fluid is a sterically stabilized colloidal dispersion of ultrafine single domain ferro or ferrimagnetic particles in a suitable carrier liquid. This colloidal system exhibits a superparamagnetic (SP) behaviour at room temperature [1] and under the action of magnetic field it develops structural ordering similar to those of crystalline solids [2] or dipolar glasses [3]. Therefore, this system can serve as a model system to study the characterization of soft condensed materials.

The time dependent behaviour of magnetization, also known as magnetic viscosity is a well-known phenomenon observed in many magnetic systems and it arises from the thermal activation of magnetization changes over energy barriers. In a magnetic fluid this barrier is usually due to anisotropy [1]. The presence of anisotropy energy hinders the relaxation of magnetic system and gives rise to magnetic viscosity. In a magnetic fluid thermal equilibrium may be reached by two different relaxation mechanisms (1) Brownian rotational diffusion, (2) Neel relaxation mechanism [1]. The dominant mode of magnetization process of particle will be that which has the shortest relaxation time. In fact, large particles relax via Brownian rotation while small particles through Neel rotation. As magnetic fluid contains a distribution of particle size both mechanism will, in general contribute to the magnetization for temperature greater than the freezing point of the carrier liquid. But in an applied magnetic field the final state would be the same irrespective of the relaxation process. Recently we have studied electron spin resonance (ESR) of $Zn_{0.1}Fe_{0.9}Fe_2O_4$ magnetic fluid between temperature 77 K to 300 K and have shown the effect of these two relaxation mechanisms on ESR linewidth and resonance field [4]. In the present paper

The anisotropy value obtained from time dependent co-efficient agrees well with that obtained from ESR results.

2. Theory

In the present system the data were obtained at 68 K, which is well below the freezing point of the carrier liquid (in the present case it is kerosene, which has the temperature of freezing at around 200 K). Therefore in the system, Neel relaxation mechanism will only contribute and this is governed by a relation [5]

$$\tau^{-1} = f_0 \exp(-\Delta E/kT), \tag{1}$$

where, ΔE is the height of the energy barrier, k and T are the Boltzmann constant and temperature, respectively. The pre-exponential factor f_0 has typical values between 10^8 to 10^{13} s^{-1}. For a system of uniaxial particles having their easy axes aligned along the field direction, the height of the energy barrier is

$$\Delta E = KV(1-h)^2. \tag{2}$$

K is the anisotropy constant and V is the volume of the particle. $h = H/H_K$ and H_K is the anisotropy field. Therefore eq. (1) can be written as

$$\tau^{-1} = f_0 \exp[-(KV/kT).(1-h)^2]. \tag{3}$$

According to Bean [6], a critical volume, V_{sp}, below which the particles are superparamagnetic depends on the experimental measurement time, t. Therefore by setting $t = \tau$, a critical volume in zero magnetic field can be written as

$$V_{sp}(0,t) = kT/K[\ln(tf_0)]. \tag{4}$$

However, in a negative field, H (i.e. initially sample is saturated for $H > 0$ and then reduced to $H < 0$), the relaxation time and critical volume are dependent on the magnitude of this field by [6]

$$V_{sp}(h,t) = [kT.\ln(tf_0)]/[K(1-h)^2]. \tag{5}$$

In a magnetic fluid there exists distribution of particle sizes, resulting in three different contributions to the magnetization in the system which can be written as [7]

$$\bar{M}(h,t) = -\int_0^y L(y,h)f(y)\,dy + \int_y^\infty f(y)dy - \int_y^y f(y)\,dy, \tag{6}$$

where y is the reduced diameter D/D_m. D and D_m are the diameter of the particle and median diameter of the system, respectively, $f(y)$ is the particle size distribution function and $L(y,h)$ is the Langevin function. M is the reduced magnetization, M/M_s, and M_s is the saturation magnetization of the fluid. The first integral in (6) represents the contribution to the magnetization from the superparamagnetic fraction of the system. The second term represents the contribution from those blocked particles which remain at positive saturation and the last term represents the contribution from

Now if t_0 is an arbitrary time at which the measurement has started, the change in magnetization over the period of time t_0 to t is

$$\Delta\bar{M}(h,t) = -2\int_{y}^{y} f(y)\,dy = \bar{M}(h,t) - \bar{M}(h,t_0). \tag{7}$$

During the time of experiment y does not vary appreciably, so one can expand the integral in (7) in power series and truncate after the first term. Thus

$$\Delta\bar{M}(h,t) = -2\left[\frac{kT}{KV_m}(1-h)^{-2}.f(y(h,t_0)).\ln(t/t_0)\right]. \tag{8}$$

In a magnetic fluid, normally the size distribution obeys log-normal distribution function given by [1]

$$f(y) = (1/\sqrt{2}\pi y\sigma)\exp[-(\ln y)^2/2\sigma^2]. \tag{9}$$

Substituting (9) in (8) and considering $t_0 = 100$ s, (8) becomes

$$\Delta\bar{M}(h,t) = -[2/(75\sqrt{2}\pi\sigma)].\ln(t/100).\exp[-(\ln y(h,t_0 = 100))^2/2\sigma^2]. \tag{10}$$

Hence, the magnetization varies linearly with $\ln t$. The co-efficient S (known as magnetic viscosity coefficient) is

$$S = [2/(75\sqrt{2}\pi.\sigma)].\exp[-(\ln y(h,t_0 = 100))^2/2\sigma^2]. \tag{11}$$

The above equation can also be written in terms of coercive force, H_c for a system with randomly oriented easy axes as [7]

$$S = S_{max}\exp[-(\ln[(H_K - H_c)/(H_K - H)]]^2.2/9\sigma^2]. \tag{12}$$

Equation (12) predicts that $S = S_{max}$ for $H = H_c$.

3. Sample preparation and measurements

3.1 *Formation conditions and preparation details*

Ultrafine particles of 10% Zn substituted magnetite were prepared by chemical co-precipitation technique. Detailed experimental procedure is discussed elsewhere [4]. Particles were coated with oleic acid and dispersed in ISOPAR-M carrier. The fluid was centrifuged at 3000 rpm for 30 min. The substitution of ions was confirmed by atomic mass absorption (Fe/Zn = 28.9).

3.2 *Structural characterization*

X-ray powder diffraction pattern was recorded on Philips diffractometer, PW 1130/90 using CuKα radiation with a graphite monochromator mounted in the diffracted beam. The particle size and its distribution were measured using GEC–AEI Cornith 275 electron microscope having a linear resolution of 10 Å.

3.3 *Magnetization measurements*

Room temperature and low temperature magnetization measurements were carried out using vibrating sample magnetometer (VSM) PAR155 fitted with Oxford Instruments CF1200 flowing gas cryostat. The measurements of time dependence were taken by saturating the sample in 10 kOe and then reducing the field to the required value. The time taken for the field to stabilize after setting the current in the magnet was measured and found to be ~ 10 s. As soon as the required field was set, the clock was started and magnetization was measured at regular intervals of time for each field.

4. Results and discussion

4.1 *Structural properties*

Figure 1 shows the X-ray diffraction pattern for the powder. Structure was found to be single phase fcc spinel with a lattice constant of 8.392 ± 0.001 Å. Measurements of the line width were made to give estimates of the particle size using the Debye–Scherrer expression. The value thus obtained was 92 Å. In order to find out the cation distribution, Ritvield analysis was done. The cation distribution thus obtained is

$$(Fe^{3+}_{1.0})^{A}\,[Fe^{2+}_{0.9}\,Zn^{2+}_{0.1}Fe^{3+}_{1.0}]^{B}O^{2-}_{4}\,. \tag{13}$$

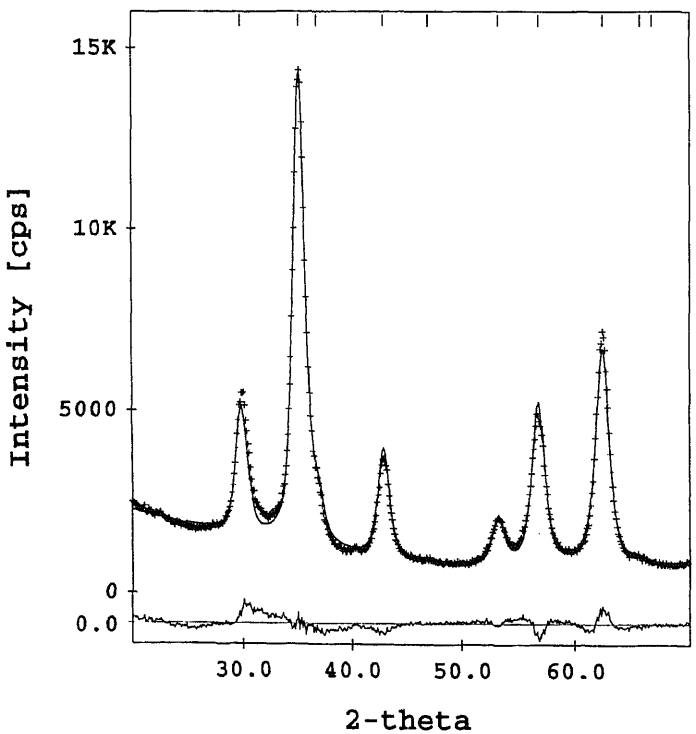


Figure 1. X-ray diffraction pattern of Fe-Zn sample.

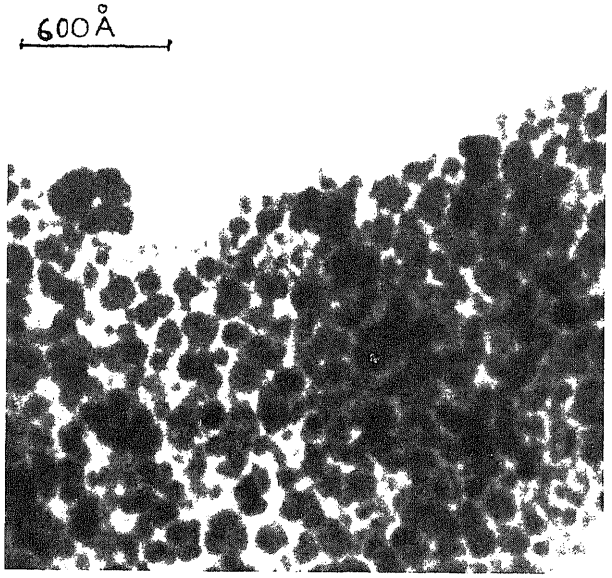


Figure 2. Electron microscope photograph for Fe-Zn sample.

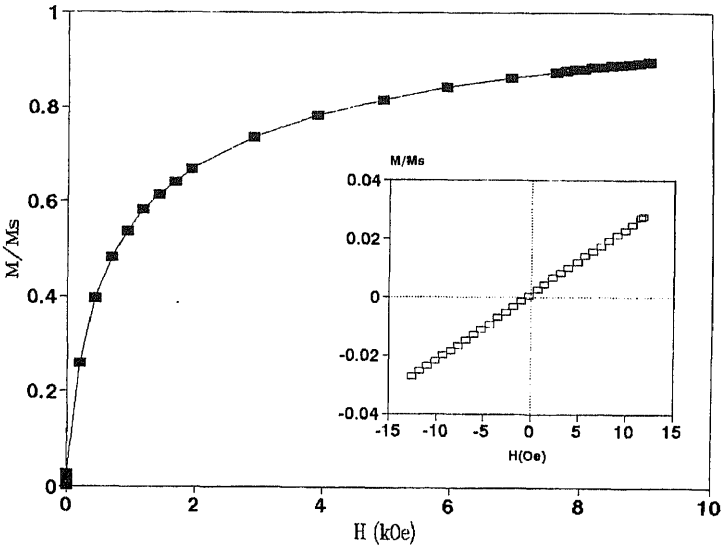


Figure 3. Room temperature reduced magnetization curve for the Fe-Zn sample. Inset: Low magnetic field data.

This result is in contradiction with the earlier results obtained for the bulk sample, where Zn occupies only tetrahedral A-site. But it has been reported that in the case of the sample synthesized by co-precipitation technique, Zn can occupy both the tetrahedral A-site and octahedral B-site of the spinel structure [8].

Figure 2 shows the electron microscope for the system. The particle diameter distribution obeys log-normal size distribution with a median diameter $D_m = 90$ Å and the standard deviation of reduced diameter $\sigma = 0.36$[9]. This value of particle size is in good agreement with the X-ray data.

4.2 *Magnetic properties*

The magnetization measurement shows that the system is superparamagnetic at room temperature (figure 3). Using the method of Chantrell *et al* [10], the median diameter and the standard deviation of the log-normal volume distribution were obtained from the room temperature magnetization data assuming particles to be spherical. The value thus obtained were $D_m = 87$ Å and $\sigma = 0.38$. This observed size represents the magnetic core size rather than the physical size. The saturation magnetization of the particles at room temperature was 62 emu/gm and increased to 64 emu/gram at 68 K. Figure 4 shows the 68 K magnetization curve. The sample exhibits ferromagnetic behaviour at 68 K i.e. having both coercivity and remanence.

Figures 5a–5c show the magnetization of the system as a function of time, t, for different values of the field, H. The solid line through the data point is the fit to eq. (10). The co-efficient, S, obtained from the fit are plotted in figure 6. The solid line shown is a curve calculated from eq. (12) using the value of $H_c = 135$ Oe, $H_K = 550$ Oe and $\sigma = 0.22$. Figure 6 shows the good fit between theory and experiment. A maximum is observed at $H = 135$ Oe which agrees well with the observed coercivity at 68 K (130 Oe).

It is known that the anisotropy field (H_K) for uniaxial particles is $2K/I_{sb}$, where I_{sb} is the bulk saturation magnetization and K is the anisotropy constant. Therefore in order to derive the value of K one needs the bulk saturation magnetization value for the system at that temperature. One way to obtain is to calculate the I_{sb} from the cation distribution assuming that this value does not change significantly ($< 5\%$) with temperature for the given temperature range, i.e. 0 K to 68 K. This is reasonable because the Curie temperature (T_c) of this sample is much higher than room temperature.

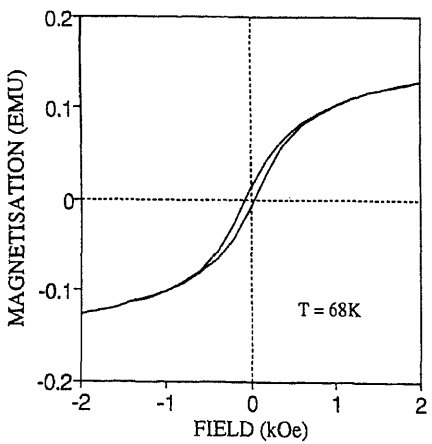


Figure 4. Magnetization versus applied field at 68 K.

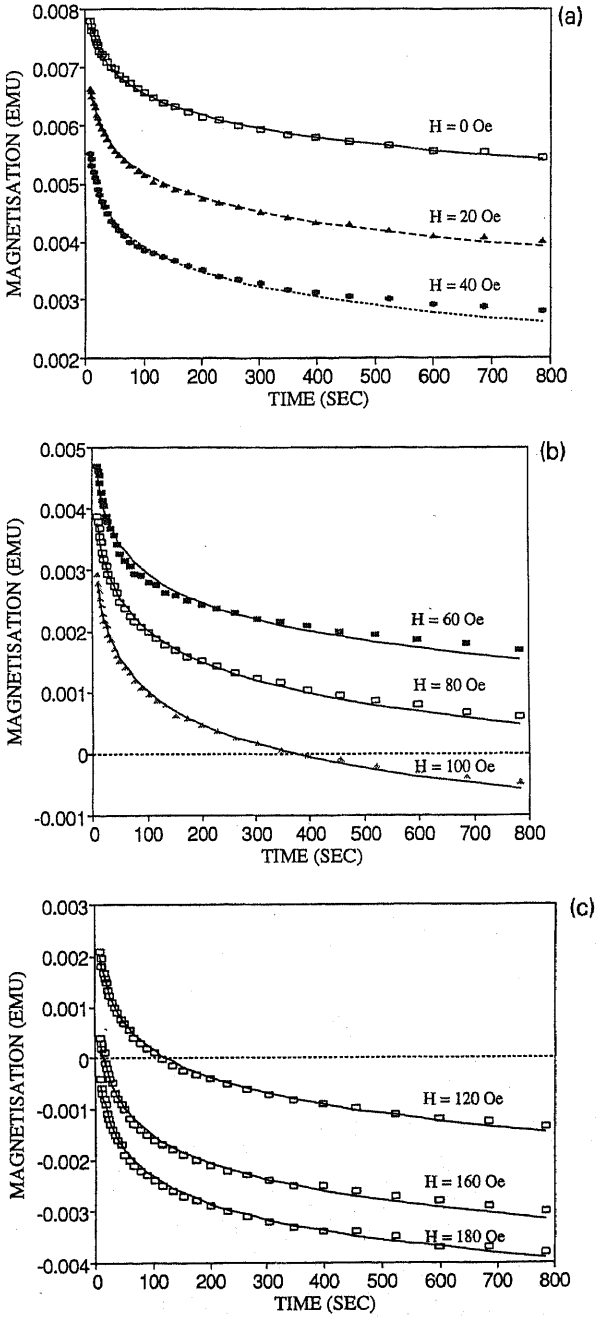


Figure 5a–5c. Time dependent magnetization curves for different field values. Line is the fit to eq. (10).

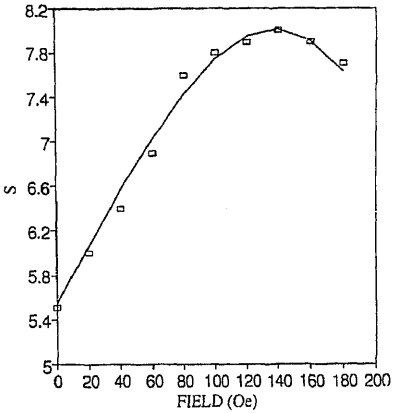


Figure 6. Magnetic viscosity co-efficient, S, versus field.

Therefore the value of T/T_c is < 0.1 at 68 K and for this I/I_{sb} is 0.98. Considering this fact and taking spin only magnetic moment of the ions, the I_{sb} for this particle was found to be 87 emu/gm. Using this value of bulk saturation magnetization, the value of K was found to be 1.2×10^5 erg/cc. This value is in agreement with the value obtained from ESR study (1.0×10^5) [9]. The observed difference in the value of magnetization of the bulk and the present sample may be due to the surface pinning effect as observed in the case of Mn substituted magnetite [11].

Acknowledgement

The work was carried out under UGC-CAREER award project of UGC. The author is thankful to Prof. R V Mehta for the comments on the paper and Dr. S W Charles, UCNW, Bangor, UK for providing experimental facilities.

References

[1] S W Charles and J Popplewell, *Ferromagnetic Materials* edited by E P Wholfarth (North Holland, Amsterdam)
[2] S D Bhagat and R V Mehta, *Indian J. Pure Appl. Phys.* **30**, 84 (1992)
[3] K O'Grady, J Popplewell and S W Charles, *J. Magn. Magn. Mater.* **39**, 56 (1983)
[4] M D Sastry, Y Babu, P S Goyal, R V Mehta, R V Upadhyay and D Srinivas, *J. Magn. Magn. Mater*, **149**, 64 (1995)
[5] L Neel, *Adv. Phys.* **4**, 191 (1955)
[6] C P Bean and J D Livingston, *J. Appl. Phys.* **30**, 120S (1959)
[7] K O'Grady, R W Chantrell, J Popplewell and S W Charles, *IEEE Trans. Mag.* **17**, 2943 (1981)
[8] B Jayadeven, K Tohji and K Nakatsuka *J. Appl. Phys.* **76**, 6325 (1994)
[9] R V Upadhyay, D Srinivas, R V Mehta and P M Trivedi, *Pramana – J. Phys.* **45**, 418 (1995)
[10] R W Chantrell, J Popplewell and S W Charles, *IEEE Trans. Mag.* **14**, 875 (1978)
[11] R V Mehta, B A Dasannacharya, P S Goyal, R V Upadhyay and K S Rao, *J. Magn. Magn. Mater*, **132**, 153 (1994)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 3,
September 1997
pp. 317–322

Fine structure in the time of flight distribution of C_2 in laser produced plasma from graphite

S S HARILAL, RIJU C ISSAC, C V BINDHU, V P N NAMPOORI and
C P G VALLABHAN
Laser Division, International School of Photonics, Cochin University of Science and Technology, Cochin 682 022, India

MS received 30 October 1996; revised 3 February 1997

Abstract. Time resolved optical emission spectroscopy is employed to study the expansion dynamics of C_2 species in a graphite plasma produced during the Nd : YAG ablation. At low laser fluences a single peak distribution with low kinetic energy is observed. At higher fluences a twin peak distribution is found. It has been noted that these double peak time of flight distribution splits into a triple peak structure at distances ≥ 17 mm from the target surface. The reason for the occurrence of multiple peak is due to different formation mechanisms of C_2 species.

Keywords. Laser produced plasma; graphite; time resolved spectroscopy; carbon clusters; recombination.

PACS Nos 52.25; 52.50; 52.70

1. Introduction

In the past few years there has been growing interest in the laser ablation of graphite target since the first report on the existence of stable carbon clusters called Buckminster fullerenes [1–6]. It has been found that when graphite target is vapourized by intense laser pulses in a helium ambient atmosphere of moderate pressures, remarkably stable carbon clusters are produced. It is expected that fullerenes are formed when vaporized carbon condenses in an atmosphere of inert gas. Although considerable progress has a been achieved in studying carbon clusters (C_n) with $n \geq 10$, little effort has been spent for the characterization and production of low mass carbon clusters C_n with $n \leq 10$. Pulsed laser deposition is increasingly being utilized for thin film preparation, in which a laser beam is focused onto the target and the resulting vapor is made to condense on a substrate of interest [7]. Diamond-like carbon thin films (DLC) which are successfully prepared using laser ablation of graphite target have wide applications owing to their interesting characteristics like extreme hardness, electrical insulation, optical transparency and high thermal conductivity [8, 9].

In order to understand the processes leading to carbon cluster formation and pulsed laser deposition, the mechanism of plasma generation under laser irradiation and the interaction of the plume with the ambient background in the plasma chamber should be investigated in detail and this has been the subject of many experimental and theoretical

investigations [10–16]. In spite of the large number of, reports available in literature, studies on laser produced carbon plasma have not yet yielded a clear cut picture on plasma dynamics of the cluster formation and such a situation arises mainly due to the complexity of the phenomenon involved. In this paper we report a comprehensive study of spatially resolved emission from C_2 molecules using time resolved spectroscopy. The temporally and spatially resolved spectroscopic studies are helpful to optimize parameters of DLC film deposition and to correlate the carbon cluster dynamics. To the best of our knowledge this is the first report regarding the existence of a triple peak in the temporal history of C_2 species from laser produced graphite plasma using time resolved emission spectroscopy.

2. Experimental setup

The experimental setup for the present study has been described elsewhere [17, 18]. A Q-switched Nd: YAG laser with fundamental frequency (1.06 μm, pulse width 9 ns) with repetition frequency 10 Hz is focused on a graphite target using a convex lens ($f = 23$ mm). The target is placed in a chamber with quartz windows in which helium gas at a pressure of 100 mTorr is maintained. The apparent length of the plasma at these pressures is about 25 mm. To avoid errors due to local heating and drilling, the sample is rotated about an axis parallel to the laser beam. The emission from the laser produced carbon plasma is imaged to produce a 1:1 image at, the entrance slit of a high resolution 1 m Spex monochromator, using appropriate focusing and collimating lenses and apertures, allowing spatial resolution better than 0.2 mm. The monochromator is equipped with a Thorn EMI Photomultiplier tube (PMT, model KQB 9863, rise time < 2 ns), connected to a digital storage oscilloscope (Iwatsu, model DS 8621, maximum sampling rate 200 MHz) with 50 ohm termination. The oscilloscope is synchronized with laser pulse using an internal trigger from the laser. The time of flight (TOF) of C_2 emission is recorded by setting the monochromator so as to record the spectral emission at $\lambda = 516.5$ nm corresponding to the (0,0) band of the Swan system ($d^3\Pi_g \rightarrow a^3\Pi_u$). Studies are made for various laser fluences as well as by varying the spatial separation normal to the target surface.

3. Results and discussion

Time resolved studies of emission lines from C_2 species were made using oscilloscope traces. Each temporal profile represents a complex convolution of different factors that governs the temporal history of the emitting species, its production mechanism and rate, its flight past the viewing point and its radiative and collisional decay rate. Time resolved emission from any species shows a definite time delay for emission with respect to the laser pulse. The time difference between the initiation of the plasma and the time taken by the species to reach a specific point inside the plasma is called time delay of that particular species. There are many reasons for the occurrence of time delay for emission from any species which includes time of flight (TOF) of the species, thermal processes leading to plasma from the target, diffusion of the species through electric field created by other charged species, recombination and/or dissociation of the species etc. For e.g., if

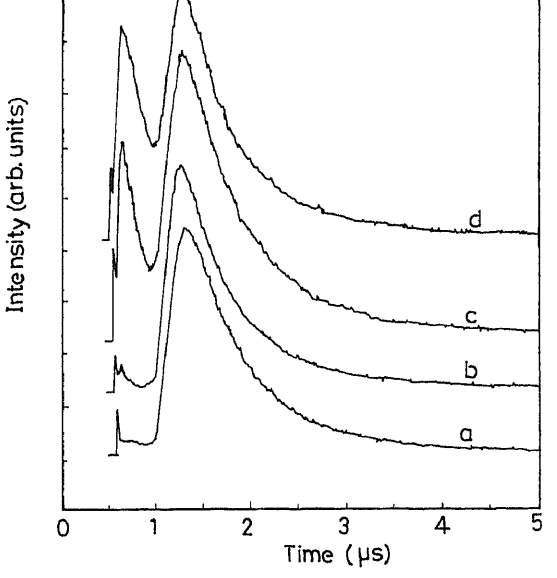


Figure 1. Intensity variation with time for 516.5 nm spectral emission from C_2 observed at 5 mm away from the target for laser fluences (**a**) 12.7 J cm^{-2} (**b**) 26.7 J cm^{-2}, (**c**) 28 J cm^{-2} and (**d**) 29.3 J cm^{-2}.

optical emission arises from electron impact excitation of species in the laser produced plasma, the excitation rate is proportional to the electron and species densities. Thus the time dependence of the emission intensity would be affected by the motion of electrons and that of the particles. Recently we reported spatial and temporal studies of C_2 species which reveal a double peak structure in time of flight (TOF) emission pattern beyond a threshold laser fluence [19]. Below this threshold laser fluence, only single peak TOF distribution is observed. The faster peak keeps on narrowing and the slower one becomes broadened with increasing helium pressure [20]. It is also noted that intensities of these peaks vary with laser fluence and helium gas pressure. The typical temporal profile for C_2 species (choosing $\lambda = 516.5$ nm corresponding to (0,0) band of Swan system at a distance 5 mm from the target for different laser fluences are given in figure 1. The initial spike in the figure is due to scattering caused by the laser beam, and can be used as a time marker. These double peak structure of C_2 species from carbon plasma were assigned due to species corresponding to those generated by dissociation of higher clusters which gives rise to slower velocity components, while those generated by many body recombination apparently gives rise to faster peak which are present only at higher laser fluences [19].

The present paper reports the observation of a triple peak structure in the temporal profile of C_2 species at distances ≥ 17 mm from the target surface. TOF spectra for C_2 species at various distances from the target is given in figure 2(a–f) for a laser fluence of 32 J/cm^2. For distances less than 17 mm from the target surface, there exist only a double peak structure in the TOF distribution (figures 2(a) and (b)). Thus, there exists a critical distance below which only double peak structure is observed while triple fold structure is

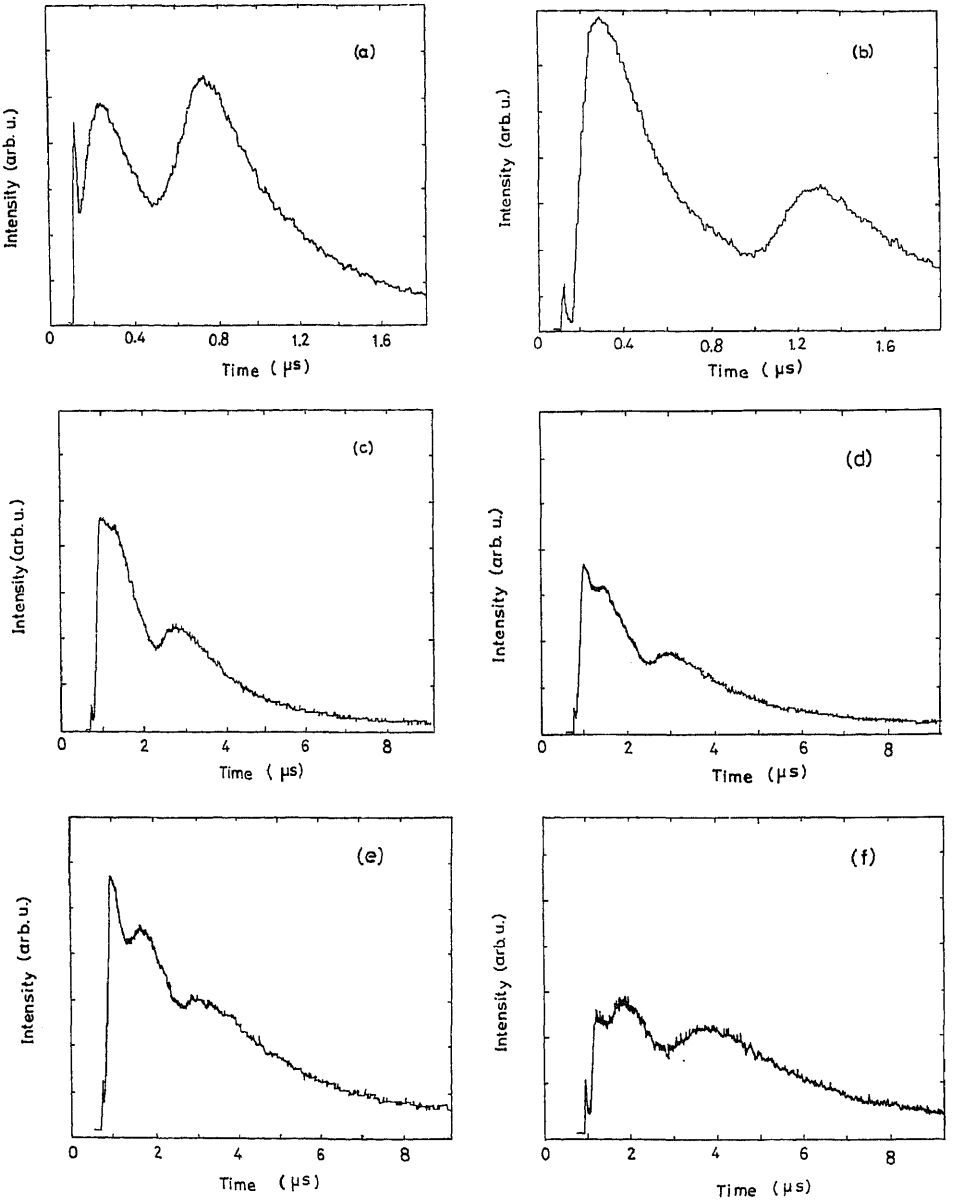


Figure 2. Intensity variation of spectral emission with time for C_2 species (516.5 nm) at different distances from the target. Distances are (**a**) 4 mm, (**b**) 10 mm, (**c**) 17 mm, (**d**) 18 mm, (**e**) 20 mm and (**f**) 22 mm. These TOF spectra are recorded at a laser fluence of 32 J/cm^2 and at a helium pressure of 100 mTorr.

observed beyond that critical distance. There are few reports which describe the triple fold time of flight (TOF) distribution in laser generated plasma from graphite target. Lowndes *et al* [21] recently observed three 'modes' of incident species in the TOF profile

using ion probe method and they attributed it to scatterd ions, ions that are slowed by gas phase collisions and slow moving clusters formed through collisions respectively. Tasaka *et al* [22,23] observed triple fold plume structure during Nd:YAG laser ablation of graphite in helium ambient atmosphere and during optical emission studies using streak camera. They found that the fastest component is composed of carbon ions, second fastest component is due to compressed neutral molecules and the slowest component is the radial vapor from the graphite target. But the triple peak structure observed in these studies using time resolved spectroscopic technique is a highly spatial dependent one.

Two mechanisms exists for the C_2 formation, viz, dissociative and recombinational mechanisms. Most of the models of fullerene formation are based on the recombination mechanism, i.e., the formation by nucleation from carbon vapour consisting of carbon atoms and very small carbon molecules [24]. It is well known that graphite exhibits a large difference between the inter-layer and the intra-layer bond strengths. At low fluence, graphite will be ablated layer by layer producing large particles which in turn get dissociated to form C_2 species [25]. The dissociative mechanism can further be supported by the observation of long duration of Swan band emission at low fluences. At low laser fluences as in figure 1, only a slowly propagating component with low kinetic energy is observed. The larger masses of C_n will result in longer delays which are observed in the C_2 emission (dissociative mechanism) occurring at the lower fluences. The dominant mechanism for the production of C_2 Swan band emission at low fluences is the electron collision with C_n cations and neutrals ($n > 2$) followed by dissociation where one of the fragments is an ejected C_2 molecule [15]. As the laser fluence is increased, clusters with lower values of n will be ejected directly from the target. Above a threshold laser fluence, temperature of the plasma becomes so large so as to dissociate C_n to neutral and ionized carbon atoms just outside the target. During spectral analysis, we observed that at low fluence levels, the emission spectrum is dominated by C_2 molecules whereas at higher fluences the plasma emission is mainly due to ionic species of carbon up to C IV along with C_2 species. At higher fluence levels, Swan band formation is mainly due to electron–ion and ion–ion recombination. It is observed here that at high laser fluences, after a threshold, an emission peak showing a faster component with higher kinetic energy for C_2 molecules begins to appear. The observation of highly energetic peak at high laser fluences is attributed to the many body recombination of these carbon ions and electrons.

From the studies of ionic species it is found that the concentration of ionic species depend very much on the laser intensity. The expansion velocities of ionic species are found to be increased with increasing ionization. The expansion velocities calculated from the kinetic energy distribution of these ionic species are found to be maximum at 40 km/s, 58 km/s and 80 km/s for C II, C III and C IV respectively at a laser fluence of 32 J/cm^2. Hence most of the highly ionized atoms are located at the front of the expanding plasma which are followed later in time by species of lower stages of ionization. It is noted that the newly formed peaks have almost identical kinetic energy distributions in comparison with highly energetic ions. The observation of splitting of the recombinational peak at higher spatial distance is due to delays caused by different recombinational formation mechanism of C_2 species in the laser induced carbon plasma. The presence of helium ambient gas helps cooling of plasma so as to reduce the plasma temperature and confinement of the same, thereby increasing the recombination rate.

In conclusion, we have observed a triple peak structure in the time profile of C_2 species

in laser induced graphite plasma. We observed a double peaks, structure in the temporal profile at higher laser fluences and these double peak structure gets modified into triple peak distribution at distances ≥ 17 mm from the target. The cause of formation of these triple peak are assigned to the delay caused by different formation mechanism of C_2 species.

Acknowledgements

The present work is supported by the Department of Science and Technology, Government of India. One of the authors (SSH) is grateful to CSIR, New Delhi for a research fellowship. The authors (RCI) and (CVB) are thankful to UGC, New Delhi for their research fellowships.

References

[1] H W Kroto, R J Heath, S C O'Brien, R F Curl and R E Smalley, *Nature* (*London*) **318**, 165 (1985)
[2] E A Rohlfing, *J. Chem. Phys.* **93**, 7851 (1990)
[3] G Meijer and D S Bethane, *J. Chem. Phys.* **93**, 7800 (1990)
[4] R E Smalley, *Acc. Chem. Res.* **25**, 98 (1992)
[5] E A Rohlfing, D M Cox and A Kaldor, *J. Chem. Phys.* **81**, 3322 (1984)
[6] Jr William Weltner, R J Van Zee, *Chem Rev.* **89**, 1713 (1989)
[7] D B Chrisey and G K Hubler (eds) *Pulsed laser deposition of thin films* (John Wiley and Sons, New York, 1994)
[8] D L Pappas, K L Saenger, J Bruley, W Krakow, J J Cuomo, T Gu and R W Collins, *J. Appl. Phys.* **71**, 5675 (1992)
[9] F Davanloo, E M Juengerman, D R Jander, T J Lee and C B Collins, *J. Appl. Phys.* **67**, 2081 (1990)
[10] P Heszler, J O Carlsson and P Demirev, *Phys. Rev.* **B53**, 12541 (1996)
[11] C P Safvan, F A Rajgara, V Bhardwaj, G R Kumar and D Mathur, *Chem. Phys. Lett.* **255**, 25 (1996)
[12] G Meijer and D S Bethune, *J. Chem. Phys.* **93**, 7800 (1990)
[13] W R Creasy and J T Brenna, *J. Chem. Phys.* **92**, 2269 (1990)
[14] C Boulmer-Leborgne, J Hermann and B Dubreulil, *Plasma Sour. Sci. Tech.* **2**, 219 (1993)
[15] T P Wright, *Phys. Rev. Lett.* **28**, 268 (1972)
[16] D W Koopman, *Phys. Fluids* **14**, 1707 (1971)
[17] S S Harilal, P Radhakrishnan, V P N Nampoori and C P G Vallabhan, *Appl. Phys. Lett.* **64**, 3377 (1994)
[18] S S Harilal, R C Issac, C V Bindhu, V P N Nampoori and C P G Vallabhan, *Pramana – J. Phys.* **46**, 145 (1996)
[19] S S Harilal, R C Issac, C V Bindhu, V P N Nampoori and C P G Vallabhan, *J. Appl. Phys.* **80**, 3561 (1996)
[20] S S Harilal, R C Issac, C V Bindhu, V P N Nampoori and C P G Vallabhan, *Jpn. J. Appl. Phys.* **36**, 134 (1997)
[21] D H Lowndes, D B Geohegan, A A Puretzky, D P Norton and C M Rouleau, *Science* **273**, 898 (1996)
[22] Y Tasaka, M Tanaka and S Usami, *Jpn. J. Appl. Phys.* **34**, 1673 (1995)
[23] Y Tasaka, M Tanaka and S Usami, *Appl. Surf. Sci.* **79/80**, 141 (1994)
[24] H W Kroto, A W Allaf and S P Balm, *Chem. Rev.* **91**, 1213 (1991)
[25] Y Iida and E S Yeung, *Appl. Spect.* **48**, 945 (1994)

physics
October 1997
pp. 323-370

Some aspects of Kaluza–Klein cosmology

S K SRIVASTAVA
Department of Mathematics, North Eastern Hill University, Permanent Campus, Shillong 793 022, India

MS received 31 October 1995; revised 18 March 1997

Abstract. This article contains a brief account of Kaluza–Klein theory and its applications in cosmology from the very beginning.

Keywords. Multidimensional gravity; Kaluza–Klein theory; quantum fields in curved space-time; early universe.

PACS Nos 04.50; 04.62

1. Motivation and historical development of Kaluza–Klein theory

There exist four fundamental forces in the Nature. These are gravitation, electromagnetism, weak and strong interactions. Unification of all these forces has been the ultimate goal of a theoretical physicist. Firstly, the unification of electricity and magnetism was obtained by Maxwell in his theory of electromagnetism in 1861–62. After the advent of special theory of relativity, three-space and time were unified into a four-dimensional space-time continuum by Minkowski. Unification of electromagnetism and weak interaction was obtained by Glashow, Weinberg and Salam. In 1961, Glashow [1] unified the weak and electromagnetic interactions using the gauge group $SU(2) \otimes U(1)$. Later on, Weinberg and Salam showed independently that weak gauge bosons can acquire mass without any loss of renormalizability [2]. This theory got experimental support in 1973 followed by discovery of the weak gauge bosons in 1983. Unification of strong, weak and electromagnetic interactions had been discussed by Georgi and Glashow, in 1974, based on the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ in the standard model of grand unified theories [3]. Various other gauge groups also have been used to obtain a satisfactory version of the grand unified theory.

In all these attempts, gravity (the oldest known interaction of the Nature) has been kept quite away, which is the source of dissatisfaction for a theoretical physicist. In 1914, Nordström [4] proposed to unify Maxwell's theory with his theory of gravity induced by scalar fields in a very imaginative way in the framework of 5-dimensional theory. In 1919, Kaluza proposed Einstein type theory of gravity in 5-dimensional space-time where 4-dimensional Einstein's gravity and electromagnetism are obtained using the cylindrical condition which implies independence of components of metric tensor from extra dimensional coordinates. In the beginning, Einstein did not like Kaluza's idea, but eventually he supported it and this work was published in 1921 [5]. In 1926, Klein found

that Kaluza's results could be obtained using quantum mechanics also [6]. For a long time, Kaluza–Klein theory was almost idle barring few papers here and there. In the middle of 1970s, string theory [7] and supergravity [8] appeared in the arena of field theory with high hopes to unify gravity and other fundamental forces. It was found that the string theory also leads to a higher-dimensional geometry like Kaluza–Klein theory. So the importance of a higher-dimensional theory was realized in true sense. In this context, Kaluza–Klein theory became very much popular by the end of 1970s. In 1978, Cremmer, Julia and Scherk proposed 11-dimensional theory which could yield gravity and the standard model of grand unified theory using Kaluza's mechanism [9]. This model of 11-dimensional theory is called $N = 1$ supergravity. After this work, people started to think that the only way to do supergravity is via Kaluza–Klein theory. Later on, 10-dimensional theory for supergravity was also proposed to cure some problems of 11-dimensional supergravity , but it contained some other degrees of freedom of Yang–Mills multiplets in contrast to 11–dimensional theory. Currently, this theory has been abandoned, but Kaluza–Klein is still interesting in the context of the early universe at high energy level where gravity is supposed to be unified with all other interactions. In what follows, a brief account of latest developments of the theory and mathematical techniques are discussed.

2. Compactification

Our observable universe is 4-dimensional. So it seems difficult to accept the idea of higher dimensional geometry on a physical ground. A solution to this problem, namely 'spontaneous compactification', added a new dimension to the Kaluza–Klein theory. According to this idea, the higher dimensional geometry is supposed to be described as a product manifold $M^4 \otimes M^K$, where M^4 is the usual 4-dimensional space-time and M^K is a compact manifold [10–12]. At extremely high energy, the universe is supposed to be $(4 + K)$-dimensional, but below a certain energy level extra-dimensional manifold M^K gets compactified to extremely small size to be resolved. As a result, only M^4 is observed. Planck length is supposed to be a possible scale of compactification. The compact manifolds are homeomorphic (topologically equivalent) to a connected sum of tori or a sphere or a projective plane [13]. Physically, it means that M^K can be conveniently deformed to one of these manifolds. Thus, for all physical purposes, a compact manifold can be treated as a connected sum of tori or a sphere or a projective plane.

3.

(a) *Kaluza–Klein model of electromagnetism*

Kaluza had shown that electromagnetism could be obtained from 5-dimensional general relativity. For this purpose, he used the method of weak-field approximation. Later on, Klein showed that Kaluza's result could be obtained, in general also, without this approximation.

In Kaluza–Klein's theory, the manifold is taken as product of M^4 and a circle S^1. Circle is a compact manifold with initial and final points identified. So a field $f(x, y)$

(here x represents four coordinates on M^4 and y is a coordinate on S^1) satisfies the condition [2]

$$f(x,y) = f(x,y+2\pi r), \tag{3.1}$$

where r is the radius of the circle. Fourier expansion for $f(x,y)$ can be written as

$$f(x,y) = \sum_{n=-\infty}^{\infty} f_n(x)\, e^{in(y/r)},$$

which satisfies the condition given by (3.1). Using the quantum operator $-i\hbar(\partial/\partial y)$ on $f(x,y)$, eigenvalue of the momentum operator corresponding to y coordinate is obtained as $(n\hbar/r)$ for a particular n. So y component of momentum will be very high for sufficiently small r which means that only $n=0$ mode can appear in the low energy physics. Due to this reason, Klein took all observed states independent of y which is analogous to Kaluza's cylindrical condition.

In Kaluza's model, the 5-dimensional line-element is taken as [14]

$$dS^2 = g_{\mu\nu}\, dx^\mu dx^\nu - \phi^2(x)[\kappa A_\mu\, dx^\mu + dy][\kappa A_\nu\, dx^\nu + dy]. \tag{3.2}$$

Here $g_{\mu\nu}$ are components of the metric tensor on M^4 with $g_{\mu 4} = g_{4\mu} = -\phi^2 \kappa A_\mu$ and $g_{44} = -\phi^2$ as well as κ is a constant of length dimension keeping κA_μ dimensionless. To discuss general relativity with this line-element is a bit difficult due to the presence of components of electromagnetic field in the metric tensor components. So, for convenience, it is better to choose basis one-forms [15]

$$\hat{\theta}^\mu = dx^\mu, \tag{3.3a}$$

$$\hat{\theta}^4 = dy + \kappa A_\mu\, dx^\mu. \tag{3.3b}$$

This type of basis is called horizontal lift basis (HLB). In HLB, components of the metric tensor, obtained from the line-element (3.2), are $\hat{g}_{\mu\nu}$, $\hat{g}_{\mu 4} = g_{4\mu} = 0$ and $\hat{g}_{\mu 4} = -\phi^2$. Basis vectors dual to bases $\hat{\theta}^{\hat{\mu}}(\hat{\mu} = 0,1,2,3,4)$ are given as

$$\hat{e}_\mu = \frac{\partial}{\partial x^\mu} - \kappa A_\mu \frac{\partial}{\partial y}, \tag{3.4a}$$

$$\hat{e}_4 = \frac{\partial}{\partial y}. \tag{3.4b}$$

The dual bases $\hat{e}_{\hat{\mu}}$ obey the algebra

$$[\hat{e}_{\hat{\mu}}, \hat{e}_{\hat{\nu}}] = \hat{C}^{\hat{\lambda}}_{\hat{\mu}\hat{\nu}} \hat{e}_{\hat{\lambda}}, \tag{3.5}$$

where $[\hat{e}_{\hat{\mu}}, \hat{e}_{\hat{\nu}}]$ are commutators of $\hat{e}_{\hat{\mu}}$ and $\hat{e}_{\hat{\nu}}$. Using the definition of $\hat{e}_{\hat{\mu}}$, given by (3.4), one finds that structure constants

$$C^4_{\mu\nu} = C_{\mu\nu 4} = -\kappa F_{\mu\nu} \tag{3.6}$$

and all other structure constants vanish. It shows that bases $\hat{e}_{\hat{\mu}}$ are anholonomic [15] (all structure constants vanish for holonomic bases). The connection coefficients in anholonomic basis is defined as

$$\hat{\Gamma}^{\hat{\mu}}_{\hat{\nu}\hat{\lambda}} = \tfrac{1}{2}\hat{g}^{\hat{\mu}\hat{\alpha}}[\hat{e}_{\hat{\lambda}}(\hat{g}_{\hat{\alpha}\hat{\nu}}) + \hat{e}_{\hat{\nu}}(\hat{g}_{\hat{\alpha}\hat{\lambda}}) - \hat{e}_{\hat{\alpha}}(\hat{g}_{\hat{\lambda}\hat{\nu}}) + \hat{C}_{\hat{\alpha}\hat{\nu}\hat{\lambda}} + \hat{C}_{\hat{\alpha}\hat{\lambda}\hat{\nu}} - \hat{C}_{\hat{\lambda}\hat{\nu}\hat{\alpha}}]. \tag{3.7}$$

So non-vanishing connection coefficients for $\hat{\Gamma}^{\hat{\lambda}}_{\hat{\mu}\hat{\nu}}$ are

$$\hat{\Gamma}^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\nu\lambda}, \tag{3.8a}$$

$$\hat{\Gamma}^{\mu}_{\nu 4} = \hat{\Gamma}^{\mu}_{4\nu} = -\tfrac{1}{2}\kappa F^{\mu}_{\nu}, \tag{3.8b}$$

and

$$\hat{\Gamma}^{4}_{\nu\lambda} = -\hat{\Gamma}^{4}_{\lambda\nu} = \tfrac{1}{2}\kappa F_{\lambda\nu}. \tag{3.8c}$$

Components of the Riemann tensor are given as

$$\hat{R}^{\hat{\mu}}_{\hat{\nu}\hat{\alpha}\hat{\beta}} = \hat{e}_{\hat{\alpha}}(\hat{\Gamma}^{\hat{\mu}}_{\hat{\nu}\hat{\beta}}) - \hat{e}_{\hat{\beta}}(\hat{\Gamma}^{\hat{\mu}}_{\hat{\nu}\hat{\alpha}}) + \hat{\Gamma}^{\hat{\mu}}_{\hat{\sigma}\hat{\alpha}}\hat{\Gamma}^{\hat{\sigma}}_{\hat{\nu}\hat{\beta}} - \hat{\Gamma}^{\hat{\mu}}_{\hat{\sigma}\hat{\beta}}\hat{\Gamma}^{\hat{\sigma}}_{\hat{\nu}\hat{\alpha}} - \hat{\Gamma}^{\hat{\mu}}_{\hat{\nu}\hat{\sigma}}\hat{C}^{\hat{\sigma}}_{\hat{\alpha}\hat{\beta}} \tag{3.9}$$

with $\hat{R}_{\hat{\nu}\hat{\beta}} = \hat{R}^{\hat{\alpha}}_{\hat{\nu}\hat{\alpha}\hat{\beta}}$ and $\hat{R} = \hat{g}^{\hat{\nu}\hat{\beta}}\hat{R}_{\hat{\nu}\hat{\beta}}$.

Ricci scalar for the line-element, given by (3.2) is calculated to be

$$\hat{R} = R - \frac{1}{4}\phi^2 F_{\mu\nu}F^{\mu\nu} - \frac{2}{3\phi^2}\phi_{,\mu}\phi_{,\nu}g^{\mu\nu}, \tag{3.10}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The volume element is

$$\sqrt{|\hat{g}|}\,\hat{\theta}^0 \wedge \hat{\theta}^1 \wedge \hat{\theta}^2 \wedge \hat{\theta}^3 \wedge \hat{\theta}^4 = \sqrt{|\hat{g}|}\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}y$$

as $\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = -\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu$.

The Einstein–Hilbert action in 5-dimensional Einstein's theory is written as

$$S = -\frac{1}{16\pi G_5}\int \mathrm{d}^4x\,\mathrm{d}y\phi\sqrt{|g|}\left[R - \frac{1}{4}\phi^2 F_{\mu\nu}F^{\mu\nu} - \frac{2}{3\phi^2}\phi_{,\mu}\phi_{,\nu}g^{\mu\nu}\right]. \tag{3.11}$$

If the radius of the compact component of $M^4 \otimes S^1$ is r,

$$0 \le y \le 2\pi r. \tag{3.12}$$

Thus,

$$S = -\frac{1}{16\pi G_4}\int \mathrm{d}^4x\,\phi\sqrt{|g|}\left[R - \frac{1}{4}\phi^2 F_{\mu\nu}F^{\mu\nu} - \frac{2}{3\phi^2}\phi_{,\mu}\phi_{,\nu}g^{\mu\nu}\right], \tag{3.13}$$

where $G_4 = G_5/2\pi r$ is the 4-dimensional gravitational constant. In the original work of Kaluza, $\phi(x)$ is taken as a constant.

(b) *Non-abelian Kaluza–Klein theories*

In 1954, Yang and Mills generalized local $U(1)$ invariance (which is an abelian theory) to local non-abelian theories. Around the same time, Shaw also did this kind of work but his contribution was not found interesting by physicists. The Yang–Mills theory was almost idle for a long span of time. In the early 1970s, this theory blossomed in the context of gauge theories, as gauge theorists found it a very convenient mathematical tool. In 1963, De-Witt proposed unification of Yang–Mills fields with gravitation under Kaluza–Klein scheme in a higher dimensional geometry [16]. A detailed discussion of this idea appeared in the work of Kerner [17]. $(4+N)$-dimensional Einstein–Hilbert

action as well as derivation of gravitation, Yang–Mills fields and scalar field was discussed by Cho and Freund in 1975 [18].

In 1982, Salam and Strathdee raised a very important question 'What are the implications of extra dimensions in our observable universe?' In [19], they have discussed that through expansion of physical fields in normal modes associated symmetry groups in Kaluza–Klein theory make their appearance as massive multiplets and through components of the metric tensor these groups yield non-abelian fields. In what follows, a brief discussion on these aspects is given.

In the $(4+N)$-dimensional theory, coordinates are separated into $x^\mu(\mu = 0,1,2,3)$ of usual space-time and y^n of N-dimensional compact manifold. The components of the metric tensor are given as [20]

$$\begin{aligned}
\bar{g}_{\mu\nu}(x,y) &= g_{\mu\nu}(x) + \tilde{g}_{mn}\,\xi^n_\alpha(y)\xi^m_\beta(y)A^\alpha_\mu(x)A^\beta_\nu(x),\\
\bar{g}_{\mu n}(x,y) &= \bar{g}_{n\mu}(x,y) = -\tilde{g}_{mn}\,\xi^m_\alpha(y)A^\alpha_\mu(x),\\
\tilde{g}_{mn}(x,y) &= \tilde{g}_{mn}(y),
\end{aligned} \tag{3.14}$$

where $\xi^n_\alpha(y)$ are Killing vectors satisfying the algebra

$$\xi^m_\beta\partial_m\xi^n_\gamma - \xi^m_\gamma\partial_m\xi^n_\beta = f^\alpha_{\beta\gamma}\,\xi^n_\alpha \tag{3.15}$$

with $f^\alpha_{\beta\gamma}$ as structure constants.

The Einstein–Hilbert action in $(4+N)$-dimensional theory is written as

$$S_{(4+N)g} = -\frac{1}{16\pi G_{(4+N)}}\int \mathrm{d}^4x\,\mathrm{d}^ny\sqrt{|g_{(4+N)}|}\,R_{(4+N)} \tag{3.16a}$$

with

$$R_{(4+N)} = R_4 + R_N(y) + \tfrac{1}{4}\tilde{g}_{mn}\xi^n_\alpha(y)\xi^m_\beta(y)F^\alpha_{\mu\nu}(x)F^{\beta\mu\nu} \tag{3.16b}$$

which can be calculated in HLB as earlier. In 5-dimensional case, $N = 1$ and the compact manifold is S^1 for which Ricci scalar vanishes. In (3.16b),

$$F^\alpha_{\mu\nu} = \partial_\mu A^\alpha_\nu - \partial_\nu A^\alpha_\mu + f^\alpha_{\beta\gamma}A^\beta_\mu A^\gamma_\nu. \tag{3.17}$$

Here α,β and γ are gauge indices for which the degree of freedom is equal to the number of Killing vectors. The number of Killing vectors depend upon the Lie group on the compact manifold. For example, if the compact manifold is S^N, the symmetry group on S^N is $SO(N+1)$. The number of generators of $SO(N+1)$ is $[N(N+1)]/2$, which is the number of Killing vectors on S^N. Also the maximum number of possible Killing vectors on N-dimensional manifold is $[N(N+1)]/2$. But it is not possible to get the maximum number of Killing vectors for every compact N-dimensional manifold, because the number of Killing vectors depend on the symmetry group. The symmetry group on N-dimensional torus is $[U(1)]^N$. So possible number of Killing vectors is N, which is the number of generators on T^N. Thus T^N has less than maximum number of Killing vectors, whereas S^N accommodates maximum number of Killing vectors on a N-dimensional manifold. So S^N is maximally symmetric, but T^N is not.

4. Physical fields on $M^4 \otimes M^K$

Here we will consider compact manifolds T^K and S^K. Calculations on projective planes can be done in a similar manner.

(i) When $M^K = T^K$: T^K is K-times product of circles with different radii $r_1, r_2, \ldots, r_k$. So Fourier expansion of a physical field $\phi(x, y)$ can be written as

$$\phi(x,y) = \sum_{n_1 \cdots n_k = -\infty}^{\infty} \phi_{n_1 \cdots n_k}(x) \exp\left[i\left(\frac{n_1 y_1}{r_1} + \cdots + \frac{n_k y_k}{r_k}\right)\right] \tag{4.1}$$

and the line-element looks like

$$\mathrm{d}S^2 = g_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu - (r_1^2\,\mathrm{d}\theta_1^2 + \cdots + r_k^2\,\mathrm{d}\theta_k^2), \tag{4.2}$$

where $0 \le \theta_1, \ldots, \theta_k \le 2\pi$. Even other boson fields can be treated in a similar way.

(ii) When $M^K = S^K$: In this case, harmonics are a bit complicated. A massless scalar field ϕ obeys the Klein–Gordon equation [21]

$$\Box_{(4+K)}\phi = \frac{1}{\sqrt{|g_{(4+K)}|}} \frac{\partial}{\partial x^M}\left(\sqrt{|g_{(4+K)}|}\, g^{MN} \frac{\partial}{\partial x^N}\right)\phi = 0, \tag{4.3}$$

where g_{MN} are components of the metric tensor on $M^4 \otimes M^K$. Equation (4.3) can be written as

$$(\Box_4 + \Box_K)\phi = 0, \tag{4.4}$$

where $\Box_4$ is defined on M^4 and $\Box_K$ on M^K.

If λ is the eigenvalue of the operator $\Box_K$ for the scalar field ϕ,

$$\Box_K \phi = \lambda\phi. \tag{4.5}$$

To obtain λ, the set of all homogeneous harmonic polynomials $\hat{\phi}$ of degree l in Euclidean $(K+1)$-space are considered. A homogeneous polynomial of degree l can be written as

$$\hat{\phi}(x_1, \ldots, x_{(K+1)}) = c_{\alpha_1, \ldots, \alpha_l} x^{\alpha_1, \ldots, \alpha_l}, \tag{4.6}$$

where $c_{\alpha_1, \ldots, \alpha_l}$ are symmetric with each l taking $(K+1)$ values. Thus linearly independent $c_{\alpha_1, \ldots, \alpha_l}$ are ${}^{(N+l)}C_l$ in number. If $\hat{\phi}$ is a harmonic on $(K+1)$-dimensional Euclidean space

$$\nabla^2_{(K+1)}\hat{\phi} = 0, \tag{4.7a}$$

which yields

$$\delta^{\alpha_1\alpha_2} c_{\alpha_1, \alpha_2, \ldots, \alpha_l} = 0. \tag{4.7b}$$

The number of equations (4.7b) is ${}^{(N+l-2)}C_{(l-2)}$.

The $(K+1)$-dimensional Euclidean space is described by the line-element

$$\mathrm{d}S^2_{(K+1)} = \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \cdots + \mathrm{d}x^2_{(K+1)}. \tag{4.8}$$

The sphere S^K is embedded in $(K+1)$-dimensional Euclidean space. For the purpose of calculation of λ, it is convenient to write (4.8) in polar coordinates as

$$dS^2_{(K+1)} = dr^2 + r^2 k_{ij}\, d\theta^i d\theta^j, \tag{4.9}$$

where $0 \leq \theta_1, \ldots, \theta_{(K-1)} \leq \pi$ and $0 \leq \theta_K \leq 2\pi$. Now (4.7a) can be written as

$$\nabla^2_{(K+1)}\hat{\phi} = \left(r^{-K}\frac{\partial}{\partial r}\left(r^K\frac{\partial}{\partial r}\right) + r^{-2}\Box_K\right)\hat{\phi} = 0. \tag{4.10}$$

As $\hat{\phi}$ is homogeneous,

$$\hat{\phi} = r^l\phi. \tag{4.11}$$

So, from (4.10),

$$\Box_K\phi = -l(l+K-1)\phi. \tag{4.12}$$

Thus $\lambda = -l(l+K-1)$ with degeneracy,

$$d_S(K,l) = {}^{(K+l)}C_l - {}^{(K+l-2)}C_{(l-2)}, \tag{4.13}$$

where ${}^nC_r \equiv n!/(r!(n-r)!)$.

If ϕ is replaced by components of a vector, λ can be calculated as

$$\lambda = -[l^2 + (l-1)(K-1)], \quad l = 1, 2, \ldots, \tag{4.14}$$

with degeneracy

$$d_v(K,l) = (K+1)[{}^{(K+l)}C_l - {}^{(K+l)}C_{(l-2)}] - {}^{(K+l+1)}C_{(l+1)} + {}^{(K+l-3)}C_{(l-3)}. \tag{4.15}$$

If ϕ is replaced by divergenceless traceless rank-2 tensor components,

$$\lambda = -[l^2 + (K-1)(l-2)], \quad l = 2, 3, \ldots, \tag{4.16}$$

with degeneracy

$$\begin{aligned} d_T(K,l) = {} & \tfrac{1}{2}(K+1)(K+2)[{}^{(K+l)}C_l - {}^{(K+l-2)}C_{(l-2)}] \\ & - (K+1)[{}^{(K+l+1)}C_{(l+1)} - {}^{(K+l-3)}C_{(l-3)}]. \end{aligned} \tag{4.17}$$

If $T_{ij} = k_{ij}T$, λ can be calculated like scalars. The traceless tensor is of the form

$$Q_{ij}(z) = \nabla_i z_j + \nabla_j z_i - \frac{2}{K}k_{ij}(\nabla^l z_l). \tag{4.18}$$

Now two cases arise here:

(i) In case $\nabla^l z_l = 0, \lambda = -[l^2 + (K-1)l - (K+2)], l = 2, 3, \ldots,$ with degeneracy $d_v(K,l)$ given by eq. (4.15).

(ii) In case $z = \nabla\chi$ (χ is a scalar), $\lambda = -[l^2 + (N-1)l - 2N]$, $l = 2, 3, \ldots,$ with degeneracy $d_S(K,l)$ given by (4.13).

A spinor field ψ satisfies the higher-dimensional Dirac equation [22, 23]

$$[i\Gamma^M \nabla_M - m]\psi = 0, \tag{4.19}$$

where Γ^M are Dirac matrices satisfying the condition

$$\{\Gamma^M, \Gamma^N\} = 2g^{MN}. \tag{4.20}$$

Γ^M in (4.20) can be decomposed as (γ^μ, γ^m). Here

$$\gamma^\mu = e^\mu_a \tilde{\gamma}^a, \quad a = 0, 1, 2, 3 \tag{4.21a}$$

and

$$\gamma^m = e^m_{\acute{m}} \tilde{\gamma}^{\acute{m}}, \quad \acute{m} = 4, 5, \ldots, 3 + K. \tag{4.21b}$$

Here $\tilde{\Gamma}^{\acute{M}} \equiv (\tilde{\gamma}^a, \tilde{\gamma}^{\acute{m}})$ are $(4 + K)$-dimensional Dirac matrices in flat space-time. In (4.21)

$$\tilde{\gamma}^a = *\gamma^a \otimes I, \tag{4.22a}$$

$$\tilde{\gamma}^{\acute{m}} = *\gamma^5 \otimes *\gamma^{\acute{m}}, \tag{4.22b}$$

where $*\gamma^a$ are standard 4×4 Dirac matrices and $*\gamma^{\acute{m}}$ are Dirac matrices on $M^K \cdot *\gamma^{\acute{m}}$ and I (identity matrix) are $2^{K/2} \times 2^{K/2}$ matrices (when K is even) and $2^{(K-1)/2} \times 2^{(K-1)/2}$ matrices (when K is odd). $*\gamma^5$ in (4.22b) is given as

$$*\gamma^5 = i*\gamma^0 *\gamma^1 *\gamma^2 *\gamma^3 \tag{4.23a}$$

and

$$(*\gamma^5)^2 = I. \tag{4.23b}$$

A spinor field $\psi(x, y)$ can be expanded as

$$\psi(x, y) = \sum_i \psi_i(x)\chi_i(y), \tag{4.24}$$

where $\chi_i(y)$ satisfy the equation

$$i\gamma^m \nabla_m \chi_i = -m_i *\gamma^5 \chi_i. \tag{4.25}$$

Equations (4.19) and (4.25) yield

$$[i\gamma^\mu \nabla_\mu - (m + m_i *\gamma^5)]\psi_i(x) = 0. \tag{4.26}$$

The Lagrangian density corresponding to the Dirac equation is given by (4.26)

$$L = \bar{\psi}_i(x)[i\gamma^\mu \nabla_\mu - (m + m_i *\gamma^5)]\psi_i(x). \tag{4.27}$$

For convenience, chiral transformations are used for $\psi_i(x)$, which is given as

$$\psi_i(x) \to \psi_i'(x) = \exp[i\alpha *\gamma^5]\psi_i(x), \tag{4.28}$$

where $\alpha = \arctan(m_i/m)$. Under these transformations

$$\bar{\psi}_i(x) i\gamma^\mu \nabla_\mu \psi_i(x)$$

remains invariant and the mass term acquires the canonical form

$$\bar{\psi}_i'(x)(m^2 + m_i^2)^{1/2}\psi_i'(x).$$

(For details of calculations, see Appendix A).

5. A realistic Kaluza–Klein model

Weak interaction and electromagnetism were unified under the symmetry group $SU(2)\otimes U(1)$. Later on, strong interaction was also brought under this scheme through the symmetry group $SU(3) \otimes SU(2) \otimes U(1)$. Popular name of this model is the standard model. With the recent discovery of top quark (truth quark) the standard model has become more interesting. In 1981, Witten suggested unification of gravity with the standard model under Kaluza–Klein scheme [24]. In this scheme, $SU(3) \otimes SU(2) \otimes U(1)$ is the symmetry group on a compact manifold. The space of lowest dimension with symmetry group $SU(3)$ is CP^2 (complex projective space). For details of CP^2, one can refer to Appendix B. S^2 is the lowest dimensional manifold with $SU(2)$ symmetry and S^1 is the group manifold for $U(1)$. Thus the compact manifold with $SU(3) \otimes SU(2) \otimes U(1)$ can be $CP^2 \otimes S^2 \otimes S^1$ which is a 7-dimensional real space. $CP^2 \otimes S^2 \otimes S^1$ is not a unique manifold with symmetry group of the standard model, but it is more convenient.

6. Cosmological models in Kaluza–Klein theory

It has been discussed above that, under Kaluza–Klein scheme, higher-dimensional space-time is taken as product of the usual 4-dimensional paracompact manifold and a compact manifold. It is also supposed that the internal space (the compact manifold is the internal space) is too small to be observed. In the early 1980s, the validity of this idea was discussed by many authors [25–28]. In this scheme, there is a large discrepancy between scales of the external space-time (observed 4-dimensional space-time) and the internal space. Now the question arises 'How is this large discrepancy caused?' To answer this question, two kinds of attempts were made. One is the method of quantum instability of the internal space for contraction by Appelquist and Chodos [29]. They computed one-loop effective potential for a scalar field in 5-dimensional space-time and found that it tends to minus infinity when the fifth dimension shrinks to zero. Later on, Rubin and Roth [30] studied finite temperature effect in the analysis done by Appelquist and Chodos and found two kinds of instabilities. Meanwhile, Tsokos [31] had shown that inclusion of 5-dimensional fermionic degree of freedom could save the one-loop effective potential from divergence to minus infinity. The other method uses a dynamical evolution of the higher-dimensional universe as suggested by Chodos and Detweiler [32]. This method is the cosmological dynamical reduction, which was found interesting because the expansion of external space-time and contraction of the internal space were explained simultaneously. Firstly, Chodos and Detweiler (CD) proposed Kasner type 5-dimensional solution of the field equations. The line-element for this model is

$$dS^2 = dt^2 - t^{1/2}[dx^2 + dy^2 + dz^2] - t^{-1/2}\rho^2 d\theta^2, \tag{6.1}$$

which can be generalized as

$$dS^2 = dt^2 - \sum_{i=1}^{3} t^{2p_i}(dx^i)^2 - \sum_{j=4}^{3+D} t^{2p_j}(dy^j)^2 \tag{6.2a}$$

with

$$\sum_{i=1}^{3} p_i + \sum_{j=4}^{3+D} p_j = 1 = \sum_{i=1}^{3} p_i^2 + \sum_{j=4}^{3+D} p_j^2. \tag{6.2b, c}$$

Later on, higher dimensional Robertson–Walker type solutions were obtained. The generalized $(1 + d + D)$-dimensional Robertson–Walker line-element looks like [33–35]

$$dS^2 = dt^2 - R_d^2(t) g_{mn}\, dx^m dx^n - R_D^2(t) g_{MN}\, dy^M\, dy^N, \tag{6.3}$$

where $R_d(t)$ is the scale-factor associated with the external space and $R_D(t)$ is the scale-factor associated with the internal space. The topology of such a model is $R \otimes M^d \otimes B^D$, $g_{mn}(m, n = 1, 2, \ldots, d)$ are components of the metric tensor on B^D, which is the internal compact manifold. The non-vanishing components of Ricci tensor for the line-element, given by (6.3), are

$$R_{00} = -\left(d\frac{\ddot{R}_d}{R_d} + D\frac{\ddot{R}_D}{R_D}\right), \tag{6.4a}$$

$$R_{mn} = \left[\frac{2k_d}{R_d^2} + \frac{d}{dt}\left(\frac{\dot{R}_d}{R_d}\right) + \frac{\dot{R}_d}{R_d}\left(d\frac{\dot{R}_d}{R_d} + D\frac{\dot{R}_D}{R_D}\right)\right] g_{mn}, \tag{6.4b}$$

$$R_{MN} = \left[\frac{2k_D}{R_D^2} + \frac{d}{dt}\left(\frac{\dot{R}_D}{R_D}\right) + \frac{\dot{R}_D}{R_D}\left(d\frac{\dot{R}_d}{R_d} + D\frac{\dot{R}_D}{R_D}\right)\right] g_{MN}, \tag{6.4c}$$

where k_d is the spatial curvature of M^d and k_D is the spatial curvature of B^D. Sahadev has obtained solution of Einstein's equations for R_d and R_D taking different perfect fluids. In these models, R_d increases and R_D decreases with time. These solutions suffer from a problem that $R_D(t) = 0$ at a certain time, which is not correct because it implies that internal space vanishes at this particular time. It is true that internal space is too small to be observed, but it does not vanish. This problem is called 'crack of doom' singularity. In CD model [32], it occurs as $t \to \infty$. Matzner and Mezzacappa [36] have proposed 5-dimensional model, where this problem occurs at some time during the evolution of the model. Solutions obtained by Sahadev also, are not free from this problem. Physics requires $R_D(t)$ to stabilize asymptotically. Copeland and Toms (CT) realized the necessity of this constraint, but they found that all solutions are unstable if the higher dimensional manifold is a product of the flat 4-dimensional space-time and a sphere [37]. Actually CT did not solve $(4 + D)$-dimensional Einstein equations, but they solved equations derived from 4-dimensional action for gravity. Gleisser and Taylor (GT) have obtained solutions of 6-dimensional Einstein's equations satisfying the above criterion [38, 39]. Rosenbaum *et al* [40] have suggested diagramatic solutions for 5-dimensional Kaluza–Klein cosmological model and have shown that by an appropriate choice of parameters 'crack of doom' singularity can be avoided. In [41, 42], some interesting solutions were obtained where $R_D(t)$ stabilizes asymptotically. In what follows, a brief account of these models is given.

Here the topology of the $(4+D)$-dimensional space-time is $R \otimes M^3 \otimes T^D$ with the line-element [38]

$$\mathrm{d}S^2 = \mathrm{d}t^2 - R_3^2(t)\left[\frac{\mathrm{d}r^2}{1-k_3 r^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)\right] - R_D^2(t) \times \delta_{mn}(\mathrm{d}y^m - \kappa A_\mu^m \mathrm{d}x^\mu)(\mathrm{d}y^m - \kappa A_\nu^m \mathrm{d}x^{nu}), \tag{6.5}$$

where k_3 is the spatial curvature of M^3 with possible values $+1, 0, -1$ for closed, flat and open models respectively. Here κ is a constant of length dimension like (3.2). The matter field considered in this model is a scalar field ϕ. The complete action is

$$S = \int \mathrm{d}^4x\, \mathrm{d}^D y \sqrt{|g_{(4+D)}|} \times \left[-\frac{1}{16\pi G_{(4+D)}} R_{(4+D)} + \frac{1}{2}\{g^{MN}(D_M\phi)^*(D_N\phi) - m^2\phi^*\phi\}\right], \tag{6.6}$$

where $G_{4+D} = G_4 V_D$ (G_4 is proportional to Newtonian gravitational constant which is equal to M_P^{-2}, M_P is Planck mass), $V_D = (2\pi)^2 \rho_1 \rho_2 \cdots \rho_D$ ($\rho_1, \rho_2, \ldots, \rho_D$ are radii of circles which are components of T^D), $R_{4+D} = R_4 - \frac{1}{4}\kappa^2 \delta_{mn} F^m_{\mu\nu} F^{n\mu\nu}$, $F^m_{\mu\nu} = \partial_\nu A^m_\mu - \partial_\mu A^m_\nu$, $D_\mu = \nabla_\mu + \kappa A^m_\mu \nabla_m$, $D_m = \nabla_m$ (∇_μ and ∇_m are covariant derivative in curved space-time) and $g_{(4+D)}$ is the determinant of the metric tensor, given by (6.5).

The conditions $2/(\sqrt{|g_{(4+D)}|})(\delta S/\delta g^{MN}) = 0$ imply Einstein's field equations

$$R_{MN} - \frac{1}{2} g_{MN} R_{4+D} = 8\pi G_{4+D}\left[T^{(\phi)}_{MN} + \frac{\kappa^2}{16\pi G} T^{(F)}_{MN}\right], \tag{6.7a}$$

where

$$T^{(\phi)}_{MN} = (\partial_M\phi)^*(\partial_N\phi) - \tfrac{1}{2} g_{MN}[g^{RS}(\partial_R\phi)^*(\partial_S\phi) - m^2\phi^*\phi] \tag{6.7b}$$

(neglecting terms containing gauge fields) and

$$T^{(F)}_{MN} = \delta_{mn}(F^m_{ML} F^{nL}_N - \tfrac{1}{4} g_{MN} F^m_{RS} F^{nRS}). \tag{6.7c}$$

Maxwell's field equations for F_{MN} and ϕ are obtained from the action given by (6.6) as

$$\left(1/\sqrt{|g_{(4+D)}|}\right)\partial_M\left[\sqrt{|g_{(4+D)}|} F^{MN}\right] = 0 \tag{6.8}$$

and

$$(\Box + m^2)\phi = 0. \tag{6.8a}$$

A_μ are required to live on $R \otimes M^3$ only [38], hence due to homogeneity of space-time F_{MN} are taken as

$$F^{MN} = \begin{cases} (\epsilon^{MN}/\sqrt{|g|})F(t) & \text{for } M = \mu,\ N = \nu \\ 0 & \text{otherwise} \end{cases}, \tag{6.9}$$

where ϵ^{MN} are components of Levi–Civita tensor. Connecting (6.8) and (6.9)

$$F(t) = A/(R_D)^D. \tag{6.10}$$

Further $\phi(x,y)$ is decomposed as

$$\phi(x,y) = \phi_0(t,0,0) + \bar{\phi}(t,x,y) \tag{6.11}$$

such that

$$(\Box + m^2)\phi_0 = 0 \tag{6.12a}$$

and

$$(\Box + m^2)\bar{\phi} = 0. \tag{6.12b}$$

$\bar{\phi}$ is a small fluctuation, so cosmic dynamics is supposed to be governed by ϕ_0 only. Hence $T^{(\bar{\phi})}_{MN}$ can be ignored compared to $T^{(\phi_0)}_{MN}$. The WKB solution of (6.12a) is written as

$$\phi_0 \simeq \frac{e^{imt}}{\sqrt{V_D V_m a^3 b^D}}. \tag{6.13}$$

Now the Einstein's field equations for the line-element (6.5) are [38]

$$3\frac{\ddot{R}_3}{R_3} + D\frac{\ddot{R}_D}{R_D} = 8\pi G_{4+D}\left[|\dot{\phi}_0|^2 - m^2|\phi_0|^2 + \frac{2D\kappa^2A^2}{16\pi G_{4+D}(D+2)b^{2D}}\right], \tag{6.14a}$$

$$\frac{2k_3}{R_3^2 l^2} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{R}_3}{R_3}\right) + \frac{\dot{R}_3}{R_3}\left(3\frac{\dot{R}_3}{R_3} + D\frac{\dot{R}_D}{R_D}\right) = \frac{\kappa^2 DA^2}{(D+2)(R_D)^{2D}} - \frac{8\pi G_{4+D}m^2\phi_0^2 t_P^2}{(D+2)}, \tag{6.14b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{R}_D}{R_D}\right) + \frac{\dot{R}_D}{R_D}\left(3\frac{\dot{R}_3}{R_3} + D\frac{\dot{R}_D}{R_D}\right) = \frac{\kappa^2 D(D+4)A^2}{2(D+2)(R_D)^{2D}} - \frac{8\pi G_{4+D}m^2\phi_0^2 t_P^2}{(D+2)}, \tag{6.14c}$$

where l is the constant of unit magnitude and length dimension, which is used for the dimensional correction. To avoid the 'crack of doom' singularity, it is assumed that

$$R_D^2(t) = f^2 + [R_3(t)]^{-2}. \tag{6.15}$$

If $f^2 < [R_3(t)]^{-2}$, $[R_3(t)]^{-2} = 2t/lf + R_{03}^2$ is the solution in case $k_3 = -1$. Moreover, the integration constant R_{03} is given as

$$f^2R_{03}^2 = \tfrac{1}{12}\left[-6(D-2) + \sqrt{36(D-2)^2 + 24(D-3)^2 - 24(D+3)}\right].$$

If $f^2 > [R_3(t)]^{-2}$,

$$[R_3(t)]^2 = \left(\sqrt{4\alpha\gamma - \beta^2}/2\gamma f^2\right)\sinh\left[(2t/t_P)\sqrt{\gamma/6} + \delta\right] - (\beta/2\gamma f^2),$$

where

$$\alpha = \frac{k_3^2 t_P^2 A^2 D^2 (D+1)(D+2)}{4f^{2D}},$$

$$\beta = -\left[\frac{k_3^2 t_P^2 A^2 D^2 (D+1)(D+2)}{4f^{2D}} - 6t_P^2 f^2\right],$$

$$\gamma = 2k_3^2 t_P^2 A^2 D^2 (D+1)(D+2) f^{2D},$$

$$f^2 R_3^2 > \frac{(D-3)}{3} \geq 1, \quad \text{as} \quad D \geq 6$$

and

$$2\pi G_4 < VR_3^2 R_D^2 m.$$

For details of these solutions, one can refer to [41]. Exact solutions of Einstein's field equations have been obtained in refs [23, 42].

In [35], Sahdev has assumed that when temperature drops below $1/R$ (R being the radius of compact manifold), extra-dimensional space ceases to be dynamically effective and 4-dimensional universe emerges. During this phase, extremely hot universe cools down to a minimum temperature $1/R$ and reheats until it becomes effectively $(1+3)$-dimensional. He has also examined horizon and entropy problems in this framework. As a result, Sahdev obtained an explanation to the horizon problem of the standard big-bang cosmology and suggests that this question should be rephrased in terms of entropy as 'Does the entropy falling within a causal horizon equals at any stage to the entropy of the observed universe ?' He has obtained that causal horizon is almost equal to $R_3(d=3)$ and entropy within the causal horizon is a free parameter.

Some authors as Alvarez and Gavela have discussed heating effect of the internal space [43], which was more clarified by Bar and Brown [44]. Adapting this approach, largeness of the entropy in the early universe can be discussed [45]. Also Abbott, Barr and Ellis have discussed the possibility of large amount of inflation from Kaluza–Klein cosmological model [46].

7. Particle creation due to quantum effects

As mentioned above, dynamical behaviour of external and internal scale-factors were explored in [33–35] and later on by others also [36–42]. In these models, external manifold blows up and the internal one shrinks. While deriving these solutions of higher dimensional Einstein's field equations, perfect fluid energy-momentum tensor was used. Alvarez and Gavela [43] discussed entropy production due to shrinking of the internal manifold which was developed by Kolb *et al* [45]. According to these estimates, entropy production was far from satisfactory. Later on, Abbott *et al* [46] suggested that large number of dimensions (~ 40) of the manifold could be helpful to explain the entropy problem. In spite of all these attempts, satisfactory explanation for large entropy could not be obtained. So, study of quantum effects on Kaluza–Klein cosmologies was found necessary. In this approach, quantized matter field is considered in the background of classical gravity in the absence of fully renormalized quantum theory of gravity [47, 48]. Following this approach, it was expected that particle production due to quantum effects could explain sufficient amount of entropy generation. Another purpose of this study was to know the effect of particle creation on dimensional reduction [49–51]. This idea was motivated by the work of Zel'dovich [52], Zel'dovich and Starobinsky [53] as well as Hu and Parker [54] regarding isotropization of the universe in the convenient 4-dimensional theory.

In 1984, Maeda [55] found that isotropization process is rapid in 5-dimensional space-time like 4-dimensional case considered by Hu and Parker [54]. In the above mentioned paper, Maeda had not taken into account the closed behaviour of the internal space. The topology of the 5-dimensional manifold was $M^4 \otimes R^1$. Later on, he discussed the same

problem for manifolds $M^4 \otimes S^1, M^4 \otimes T^D$ and $M^4 \otimes S^7$, where internal manifolds were compact. Some details of similar investigations for other topologies are given as under [47].

The line-element for $(1+d+D)$-dimensional space-time $(R^1 \otimes M^d \otimes M^D)$ can be taken as given in (6.3), where M^d and M^D are maximally symmetric spaces. Here the aim is to study the effect of particle creation due to quantum process. ϕ is the matter field which is massless and conformal. So, the Lagrangian for ϕ is given as

$$L_\phi = \tfrac{1}{2}\sqrt{|g_{(4+D)}|}\{g^{MN}(\partial_M\phi)(\partial_N\phi) - \xi R\phi^2\}, \tag{7.1a}$$

where

$$\xi = \frac{d+D-1}{4(d+D)} \tag{7.1b}$$

and R is the Ricci scalar for $(1+d+D)$-dimensional space-time. The field equation for ϕ is obtained from the Lagrangian given by (7.1) as

$$(\Box + \xi R)\phi = 0. \tag{7.2}$$

Components of the energy-momentum tensor for ϕ are given by

$$T_{MN} = \partial_M\phi\partial_N\phi - \tfrac{1}{2}g_{MN}\partial^P\phi\partial_P\phi - \xi\phi^2(R_{MN} - \tfrac{1}{2}g_{MN}R) + \xi[\nabla_M\nabla_N\phi^2 - g_{MN}\nabla^P\nabla_P\phi^2]. \tag{7.3}$$

In the case of a conformal field, trace of the energy-momentum tensor vanishes at the classical level, but an anomalous term appears at quantum level when dimension of the manifold is even. So, to avoid the trace anamoly at the quantum level also, Maeda took $(d+D)$ even and set $T=0$ (see ch. 6 of [56] for further details).

Harmonics $h_l(x)$ and $H_L(y)$ can be introduced for M^d and M^D spaces respectively satisfying equations

$$\nabla_\mu\nabla^\mu h_l(x) = -k_l^2 h_l(x), [\mu,\nu = 0,1,\ldots,(1+d)] \tag{7.4a}$$

and

$$\nabla_m\nabla^m H_L(y) = -k_L^2 H_L(y), [m = 2+d,\ldots,(1+d+D)]. \tag{7.4b}$$

These harmonics obey orthonormal conditions

$$\int \mathrm{d}^d x\sqrt{|g_d|}h_l^*(x)h_{l'}(x) = \delta_{ll'} \tag{7.5a}$$

and

$$\int \mathrm{d}^D y\sqrt{|g_D|}H_L^*(y)H_{L'}(y) = \delta_{LL'}, \tag{7.5b}$$

where g_d is the determinant of $g_{\mu\nu}$ and g_D is the same for g_{mn}. ϕ can be expanded as

$$\phi = \tilde{R}(\eta)^{-(d+D-1)/2}\chi(\eta,x,y), \tag{7.6}$$

where

$$\tilde{R}(\eta) = [(R_d)^d (R_D)^D]^{1/(d+D)}$$

and

$$\eta = \int^t \frac{\mathrm{d}\tilde{t}}{\tilde{R}(\tilde{t})}.$$

Further $\chi(\eta, x, y)$ can be written as

$$\chi(\eta, x, y) = \sum_{l,L} [A_{l,L}\chi_{l,L}(\eta)h_l(x)H_L(y) + A^{\dagger}_{l,L}\chi^*_{l,L}(\eta)h^*_l(x)H^*_L(y)]. \tag{7.7}$$

Using ϕ, given by (7.6), equation (7.2) looks like

$$\chi'' + Q\chi - \tilde{R}^2(\nabla_\mu \nabla^\mu + \nabla_m \nabla^m)\chi - \xi\tilde{R}^2\left[\frac{R^{(d)}}{R_d^2} + \frac{R^{(D)}}{R_D^2}\right]\chi = 0, \tag{7.8}$$

where prime ($'$) denotes derivation with respect to η.

Further, connecting (7.7) and (7.8), one obtains

$$\chi'' + (\Omega^2_{lL} + Q)\chi_{lL} = 0, \tag{7.9a}$$

where

$$Q = \frac{dD(d+D+1)}{4(d+D)^2}\left(\frac{R'_d}{R_d} - \frac{R'_D}{R_D}\right)^2 \tag{7.9b}$$

and

$$\Omega^2_{lL} = \tilde{R}^2\left[\frac{\omega_l^2}{R_d^2} + \frac{\omega_L^2}{R_D^2}\right] \tag{7.9c}$$

with

$$\omega_l^2 = k_l^2 + d(d-1)k_d \tag{7.9d}$$

and

$$\omega_L^2 = K_L^2 + D(D-1)k_D \tag{7.9e}$$

as $R^{(d)} = d(d-1)k_d$ and $R^{(D)} = D(D-1)k_D$ because M^d and M^D are maximally symmetric spaces (homogeneous and isotropic space).

The scalar product of two spinless fields are given as

$$(\phi_1, \phi_2) = -i\int_\Sigma [\phi_1(x)\partial_M \phi_2^*(x) - \partial_M \phi_1^*(x)\phi_2(x)]n^M\sqrt{|g|}\,\mathrm{d}\Sigma, \tag{7.10}$$

where n^M is the future-directed vector given as $(1, 0, 0, \ldots, 0)$ orthogonal to space-like hyperspace Σ with $\mathrm{d}\Sigma$ as volume element which is equal to $\mathrm{d}^d x\,\mathrm{d}^D y$. With this definition, orthonormal conditions for ϕ_M and ϕ_N can be written as

$$(\phi_M, \phi_N) = \delta_{MN}, (\phi_M^*, \phi_N^*) = -\delta_{MN}, (\phi_M, \phi_M^*) = 0. \tag{7.11}$$

The conditions for canonical quantization are given as

$$[\phi(t,x,y),\phi(t,x',y')]=0,$$
$$[\pi(t,x,y),\pi(t,x',y')]=0,$$
$$[\phi(t,x,y),\pi(t,x',y')]=i\delta^d(x-x')\delta^D(y-y'), \tag{7.12}$$

where $\pi=\partial L/\partial(\partial_0\phi)$ and L is the Lagrangian density for ϕ.

The equal-time commutation relations (7.12) yield

$$[A_{lL},A_{l'L'}]=0=[A^\dagger_{lL},A^\dagger_{l'L'}] \tag{7.13a, b}$$

and

$$[A_{lL},A^\dagger_{l'L'}]=\delta_{ll'}\delta_{LL'} \tag{7.13c}$$

for creation and annihilation operators $A^\dagger_{lL}$ and A_{lL} respectively. The vacuum state $|0\rangle$ is defined as

$$A_{lL}|0\rangle=0 \tag{7.14}$$

for all l and L . The orthonormality condition is obtained as

$$\chi^{*\prime}_{lL}\chi'_{lL}-\chi^*_{lL}\chi'_{lL}=0. \tag{7.15}$$

Using (7.6), (7.8) and components of the metric tensor given by (6.3) in (7.3),

$$\begin{aligned}T^0_0=&\frac{1}{2}\tilde{R}^{-(d+D+1)}[\chi'^2-Q\chi^2+\tilde{R}^2\{R_d^2[(4\xi-1)(\nabla^\mu\chi)\nabla_\mu\chi\\&+4\xi\chi\nabla^\mu\nabla_\mu\chi-\xi R^{(d)}\chi^2]+R_D^2[(4\xi-1)(\nabla^m\chi)\nabla_m\chi\\&+4\xi\chi\nabla^m\nabla_m\chi-\xi R^{(D)}\chi^2]\}].\end{aligned} \tag{7.16}$$

The vacuum expectation value of T^0_0 yields the energy density as $\rho=\langle 0|T^0_0|0\rangle$ which is independent of spatial coordinates, so one can write

$$\rho=\frac{1}{V_dV_D}\int \mathrm{d}^dx\,\mathrm{d}^Dy\sqrt{|g|}\langle 0|T^0_0|0\rangle, \tag{7.17}$$

where $V_d=\int \mathrm{d}^dx\sqrt{|g|}$ and $V_D=\int \mathrm{d}^Dy\sqrt{|g|}$. Using (7.14) and (7.16) in (7.17), one can write

$$\rho=\frac{\tilde{R}^{-(d+D+1)}}{2V_dV_D}\sum_{l,L}[|\chi'_{lL}|^2+(\Omega^2_{lL}-Q)|\chi_{lL}|^2], \tag{7.18}$$

which includes ultraviolet divergence. It means that T_{MN} need regularization. The appropriate method for this case is the adiabatic regularization given by Hu and Parker [see Appendix C].

Evaluation of $\sum_l$ or $\sum_L$ depend on the topology of the space. If d-dimensional space is topologically open or flat ie, $k_d\leq 0$,

$$\sum_l\to\frac{V_d}{(2\pi)^d}\int \mathrm{d}^dk.$$

Similarly for D-dimensional internal space

$$\sum_{L} \to \frac{V_D}{(2\pi)^D} \int \mathrm{d}^D k,$$

if $k_D \leq 0$.

Looking at the expression of ρ given by (7.19), one finds that it is divergent for large modes l or L. Technically speaking, the equation for ρ contains ultraviolet divergence. So naturally it needs regularization for being physically meaningful. In the case of time-dependent space-time background which are not conformally flat, the only appropriate regularization method is adiabatic regularization. This method too has a problem that it cannot provide energy momentum explicitly as a functional of metric unless explicit form of the line-element is available. Under these circumstances, Maeda [47] has given crude approximation of higher-dimensional version of adiabatic regularization method given by Fulling *et al* [57], which is discussed below.

In this method, energy density ρ is divided into quantum part $\rho_{q(t)}$ and classical part $\rho_{c(t)}$. Quantum domain is given by the condition $\omega_{\underline{L}} = \tilde{R}^{-1}\Omega_{\underline{L}} \leq r_H^{-1}(t)$, where $\underline{L} \equiv (l, L)$ and r_H is the horizon scale given as

$$r_H(t) = \tilde{R}(t) \int^t \frac{\mathrm{d}\tilde{t}}{\tilde{R}(\tilde{t})}.$$

The classical domain satisfies the condition $\omega_{\underline{L}} > r_H^{-1}(t)$. It can be realized below that for low frequency modes $\Omega_{\underline{L}}$, probability of creation of quantum particle is high. Here, by quantum domain, one means the domain in momentum space where quantum effects such as particle creation and vacuum polarization are dominant. Other regions in momentum space, where quantum effects are suppressed are called classical domain satisfying the above mentioned condition.

Using (7.18) as well as conditions for quantum(classical) domains, $\rho_{q(t)}$ and $\rho_{c(t)}$ can be estimated. For these estimations, solution of the differential equation (7.9) is needed. To get the required solution, method of variation of parameters can be used. To apply this method, a particular form of the complementary function is required, which is obtained here through WKB approximations. Thus solution of (7.9) can be written as

$$\chi_{lL} = (2\Omega_{lL})^{-1/2}\left[\alpha_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right) + \beta_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right]. \quad (7.19)$$

From (7.19), one obtains

$$\chi_{lL} = -i\left(\frac{\Omega_{lL}}{2}\right)^{1/2}\left[\alpha_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right) + \beta_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right] \quad (7.20a)$$

imposing the condition

$$2\Omega_{lL}\left[\alpha'_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right) + \beta'_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right]$$
$$= \Omega'_{lL}\left[\alpha_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right) + \beta_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right], \quad (7.20b)$$

which is normally used in the method of variation of parameters. Connecting (7.9) and (7.20), one obtains

$$i\left(\frac{\Omega_{lL}}{2}\right)^{1/2}\left[\alpha'_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)-\beta'_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right]$$
$$=(2\Omega_{lL})^{-1/2}\left[\left(Q-\frac{i}{2}\Omega'_{lL}\right)\alpha_{lL}\exp\left(-i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right.$$
$$\left.+\left(2\Omega_{lL}^2+Q+\frac{i}{2}\Omega'_{lL}\right)\beta_{lL}\exp\left(i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right]. \tag{7.21}$$

Equations (7.20b) and (7.21) determine

$$\alpha'_{lL}=(\Omega_{lL})^{-1}\left[-\frac{i}{2}Q\alpha_{lL}+\frac{1}{2}(\Omega'_{lL}-iQ)\beta_{lL}\exp\left(2i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right] \tag{7.22a}$$

and

$$\beta'_{lL}=(\Omega_{lL})^{-1}\left[\frac{i}{2}Q\alpha_{lL}+\frac{1}{2}(\Omega'_{lL}+iQ)\alpha_{lL}\exp\left(-2i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\right]. \tag{7.22b}$$

In the quantum domain, ω_l and ω_L can be taken small. So there is no harm in setting

$$\exp\left(2i\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right)\simeq 1. \tag{7.23}$$

Using (7.23) in (7.22), one obtains first order ordinary differential equations for $(\alpha_{lL}+\beta_{lL})$ and $(\alpha_{lL}-\beta_{lL})$ yielding solutions

$$\alpha_{lL}=C_1\left[\Omega_{lL}^{1/2}-i\Omega_{lL}^{-1/2}\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right]+C_2\Omega_{lL}^{-1/2} \tag{7.24a}$$

and

$$\beta_{lL}=C_1\left[\Omega_{lL}^{1/2}+i\Omega_{lL}^{-1/2}\int_{\eta_0}^{\eta}\mathrm{d}\bar{\eta}\,\Omega_{lL}\right]-C_2\Omega_{lL}^{-1/2}, \tag{7.24b}$$

where C_1 and C_2 are complex constants. Connecting (7.15), (7.19) and (7.20), one obtains the condition

$$|\alpha_{lL}|^2-|\beta_{lL}|^2=1. \tag{7.25}$$

Equations (7.24) and (7.25) yield

$$C_1C_2^*+C_1^*C_2=\tfrac{1}{2}. \tag{7.26}$$

If the initial state is vacuum,

$$\alpha_{lL}(\eta_0)=1 \quad \text{and} \quad \beta_{lL}(\eta_0)=0.$$

Using these initial conditions in (7.24), one obtains

$$C_1 = \tfrac{1}{2}[\Omega_{lL}(\eta_0)]^{-1/2} \quad \text{and} \quad C_2 = \tfrac{1}{2}[\Omega_{lL}(\eta_0)]^{1/2}. \tag{7.27}$$

Now $\rho_{q(t)}$ is obtained as

$$\rho_{q(t)} = (2V_d V_D \tilde{R}^{(d+D+1)})^{-1} \sum_{l,L} \times \left[\frac{1}{2\Omega_{lL}(\eta_0)}\left\{\Omega_{lL}^2 - Q + \left(\int_{\eta_0}^{\eta} Q\, \mathrm{d}\bar{\eta}\right)^2\right\} + \frac{1}{2}\Omega_{lL}(\eta_0)\right], \tag{7.28}$$

using (7.19)–(7.27) in (7.18). Initially horizon scale will be very small. So the frequency will satisfy the condition $\omega_{\underline{L}} < r_H^{-1}(t)$ because $\omega_{\underline{L}} > r_H^{-1}(t)$ will lead to unphysical situation leading to $\omega_{\underline{L}}$ being almost infinite. Thus, initially one has quantum domain. So, it is reasonable to consider quantum particles not to exist initially. As a result,

$$\rho_{c(t_0)} = 0. \tag{7.29}$$

Quantum domain which satisfies the condition $\omega_{\underline{L}} < r_H^{-1}(t) = [\tilde{R}(t)\eta(t)]^{-1}$ (where $r_H(t) = \tilde{R}(t)\int^t \mathrm{d}\tilde{t}/\tilde{R}(\tilde{t}) = \tilde{R}(t)\eta(t)$) shrinks with time. So quantum particles will enter into the classical domain satisfying the condition $\omega_{\underline{L}} < r_H^{-1}(t)$ as horizon scale will increase with time ($r_H^{-1}(t)$ will decrease with time yielding $\omega_{\underline{L}} > r_H^{-1}(t)$ for some $\omega_{\underline{L}}$ satisfying $\omega_{\underline{L}} < r_H^{-1}(t)$ earlier). Thus one obtains energy density of quantum particles entering into classical domain in time interval $(\eta, \eta + \triangle\eta)$ given as

$$\delta\rho_q(t) = \rho_q(\underline{L}_m(\eta), \eta) - \rho_q(\underline{L}_m(\eta + \triangle\eta), \eta), \tag{7.30a}$$

$\underline{L}_m$ being the maximum quantum number which satisfies the condition $\omega_{\underline{L}} < r_H^{-1}(t)$. Maeda has assumed these classical particles to behave like the collision-dominated relativistic perfect fluid with isotropic pressure $P_c = \rho_c/(d+D)$. Conservation of energy momentum tensor leads to

$$\rho_c(\eta + \triangle\eta) = [R(\eta)/R(\eta + \triangle\eta)]^{1+d+D}[\rho_c(\eta) + \delta\rho_q(\eta)],$$

which can be written as

$$\begin{aligned}
\rho(\eta + \triangle\eta) &= [R(\eta)/R(\eta + \triangle\eta)]^{1+d+D}[\rho_c(\eta) + \delta\rho_q(\eta)] \\
&\quad + \rho_q(\underline{L}_m(\eta + \triangle\eta), (\eta + \triangle\eta)) \\
&= [R(\eta)/R(\eta + \triangle\eta)]^{1+d+D}\rho(\eta) + \{[R(\eta)/R(\eta + \triangle\eta)]^{1+d+D} \\
&\quad \times (\delta\rho_q(\eta) - \rho_q(\eta)) + \rho_q(\underline{L}_m(\eta + \triangle\eta), \eta) \\
&\quad + \triangle\eta\frac{\partial}{\partial\eta}\rho_q(\underline{L}_m(\eta + \triangle\eta), (\eta + \triangle\eta))\}_{\triangle\eta=0} \\
&= [R(\eta)/R(\eta + \triangle\eta)]^{1+d+D}\rho + \triangle\eta\tilde{R}^{-(1+d+D)} \\
&\quad \times \frac{\partial}{\partial\eta}\rho_q(L_m(\eta + \triangle\eta)\tilde{R}^{-(1+d+D)}(\eta + \triangle\eta))\}_{\triangle\eta=0}.
\end{aligned} \tag{7.30b}$$

Equation (7.30) leads to a very interesting result

$$(\rho\tilde{R}^{1+d+D})' = \frac{\partial}{\partial\eta}[\rho_q\tilde{R}^{1+d+D}], \tag{7.31}$$

on taking the limit $\triangle\eta \to 0$ and using the basic definition of derivatives. Equation (7.31) relates rate of creation of quantum particles to the total energy density.

The energy momentum tensor for the perfect fluid can be written as

$$T_{MN} = (\rho + p)u_M u_N - (\delta p_d + \delta' p_D)g_{MN}, \tag{7.32a}$$

where $u^M \equiv (1, 0, 0, \ldots, 0)$ are components of the velocity vector normalized to unity,

$$\delta = \begin{cases} 1 & \text{for } M, N = \mu, \nu = 0, 1, \ldots, d \\ 0 & \text{for } M, N = m, n = 1, \ldots, D \end{cases} \tag{7.32b}$$

and

$$\delta' = \begin{cases} 0 & \text{for } M, N = \mu, \nu = 0, 1, \ldots, d \\ 1 & \text{for } M, N = m, n = 1, \ldots, D. \end{cases} \tag{7.32c}$$

With the definition of energy-momentum tensor, given by (7.4), conservation equation $T^M_{N;M} = 0$ leads to

$$V^{-1}(\rho V)' + dp_d\frac{R_d'}{R_d} + Dp_D\frac{R_D'}{R_D} = 0, \tag{7.33}$$

where $V \equiv (R_d)^d(R_D)^D = \tilde{R}^{d+D}$. Trace of energy-momentum tensor is given as

$$T = \rho - dp_d - DP_D. \tag{7.34}$$

Using (7.31), eqs (7.33) and (7.34) yield

$$p_d = -\frac{(T-\rho)\beta'}{d(\beta'-\alpha')} + \frac{\tilde{R}^{-(1+d+D)}}{d(\beta'-\alpha')}\left[\frac{\partial}{\partial\eta}[\rho_q\tilde{R}^{1+d+D}]\right], \tag{7.35a}$$

and

$$p_D = \frac{(T-\rho)\beta'}{D(\beta'-\alpha')} - \frac{\tilde{R}^{-(1+d+D)}}{D(\beta'-\alpha')}\left[\frac{\partial}{\partial\eta}[\rho_q\tilde{R}^{1+d+D}]\right], \tag{7.35b}$$

where $\alpha' = R_d'/R_d$ and $\beta' = R_D'/R_D$. Thus, rate of creation of quantum particles contribute back reaction effect on the background.

In the background model $M^4 \otimes M^D$ (where M^4 is the conventional four-dimensional flat universe and M^D is the closed D-dimensional internal space), energy density of quantum particles can be calculated at a particular time $t = t_0$ from (7.28) as

$$\rho_{q(t_0)} = [16\pi^3 V_D\tilde{R}^{4+D}]^{-1}\sum_L\int \mathrm{d}^3k\left[\frac{1}{2\Omega_L(\eta_0)}\{\Omega_L^2(\eta_0) - Q\} + \frac{1}{2}\int \mathrm{d}^3k\Omega_L(\eta_0)\right]. \tag{7.36}$$

Here, $\sum_L \to (V_3/(2\pi)^3)\int \mathrm{d}^3k$ has been used as $k_3 = 0$. Using the definition of Ω^2 from

(7.9) in (7.36), $\rho_{q(t_0)}$ is calculated as

$$\rho_{q(t_0)} = [16\pi^3 V_D \tilde{R}^{4+D}]^{-1}\left[(R_0/a_0)^2\sum_L I_2 + (R_0/b_0)^2\sum_L(\omega_L^2 I_0) - \frac{Q}{2}\sum_L I_0\right], \tag{7.37a}$$

where $a_0 \equiv a(\eta_0)$, $b_0 \equiv b(\eta_0)$, $R_0 \equiv R(\eta_0)$,

$$I_0 \equiv \int \mathrm{d}^3k[\Omega_L(\eta_0)]^{-1} = 2\pi\frac{a_0}{R_0}\left[k_m\left(k_m^2 + \frac{a_0^2}{b_0^2}\omega_L^2\right)^{1/2}\right.$$
$$\left. - \frac{a_0^2}{b_0^2}\omega_L^2 \ln\left\{\frac{k_m + [k_m^2 + (a_0\omega_0/b_0)^2]^{1/2}}{(a_0\omega_0/b_0)}\right\}\right] \tag{7.37b}$$

and

$$I_2 \equiv \int \mathrm{d}^3k[\Omega_L(\eta_0)]^{-1}k^2 = \frac{\pi a_0}{2R_0}\left[k_m[k_m^2 + (a_0\omega_0/b_0)^2]^{1/2}\left(2k_m^2 - 3\frac{a_0^2}{b_0^2}\omega_L^2\right)\right.$$
$$\left. + 3\frac{a_0^2}{b_0^2}\omega_L^2 \ln\left\{\frac{k_m + [k_m^2 + (a_0\omega_0/b_0)^2]^{1/2}}{(a_0\omega_0/b_0)}\right\}\right]. \tag{7.37c}$$

Here $k_m = a\{r_H^{-2} - (\omega_L^2/b^2)\}^{1/2}$ can be attained as given below. Using the condition for quantum domain, one obtains

$$\frac{k^2}{a_0^2} + \frac{\omega_L^2}{b_0^2} \leq r_H^{-2}(\eta_0)$$

(as M^4 is flat), which yields

$$|k| \leq a_0\{r_H^{-2}(\eta_0) - \omega_L^2/b^2)\}^{1/2}.$$

k_m is the maximum of $|k|$, so one gets the above mentioned value of the same which means

$$-k_m \leq k \leq k_m.$$

The adiabatic regularization method given by Fulling *et al* [57] is discussed in Appendix C. Using this method for the higher-dimensional case, regularized energy density can be written as

$$\rho \equiv -\langle 0_A|T_0^0|0_A\rangle = [2V_d V_D \tilde{R}^{(1+d+D)}]^{-1}\sum_l\sum_L$$
$$\times\left\{|\chi_{lL}'|^2 + (\Omega^2 - Q)|\chi_{lL}|^2 - \Omega - \frac{1}{2\Omega}\left[\left(\frac{\Omega'}{2\Omega}\right)^2 - Q\right]\right.$$
$$\left. + \frac{1}{4\Omega}\left[\left(\frac{\Omega'}{2\Omega}\right)^2\right]\epsilon_{2(2)} - \frac{\Omega'}{4\Omega}\epsilon_{2(3)}' - \frac{1}{4}\Omega^2(\epsilon_{2(2)})^2 - \frac{1}{2}Q\epsilon_{2(2)}\right\}, \tag{7.38}$$

where $\epsilon_{m(n)}$ is defined in Appendix C.

Assuming $d = 3$ and M^4 as a flat space-time

$$\rho_{q(t)} = [16\pi^3 V_D \tilde{R}^{(4+D)}]^{-1} \sum_L \int \mathrm{d}^3k \times \left\{ \frac{1}{2\Omega_0} \left[\Omega^2 - Q\left(\int_{\eta_0}^{\eta} Q\, \mathrm{d}\eta' \right)^2 \right] + \frac{\Omega_0}{2} - \Omega + (Q/2\Omega) \right\}, \tag{7.39}$$

which is the regularized energy density in the quantum region.

8. Gravitational Casimir energy in Kaluza–Klein theories

As discussed above, one gets convinced that extra dimensions should be extremely small as these dimensions are not observed at low energy level. To understand it in the case of higher dimensional static models, one can make an analogy to the Casimir effect in electrodynamics which is a quantum effect. According to classical electrodynamics, there exists no force between two conducting uncharged infinite plates. However, results, obtained from quantum electrodynamics, are different. These results yield an attractive force between such conducting plates due to negative zero-point energy between them [56, 58, 59]. In Kaluza–Klein theories, similar situation is obtained between external and internal manifolds. For example, in the simplest case of 5-dimensional theory, $x^5 = 0$ and $x^5 = L_5$ where L_5 is the distance around the fifth dimension. Appelquist and Chodos have obtained the effective potential of the metric field up to one-loop quantum correction given as [29]

$$V_{\mathrm{eff}}(L_5) = -\frac{15\zeta(5)}{4\pi^2 (L_5)^4},$$

where Riemann zeta function $\zeta(5) = 1.037\cdots$. The negative sign for V_{eff} shows that the effective potential is attractive causing contraction of the extra dimension. This result is reliable only when L_5 is greater than L_P as loop expansion is valid up to this scale only.

In what follows, the mathematical method for calculation of effective potential (Casimir energy) for gravity in $(d+D)$-dimensional space-time $M^d \otimes M^D$ is discussed [21]. The action is taken as

$$S[g] = -\frac{1}{16\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|}(R + 2\Lambda). \tag{8.1}$$

Under transformations $g_{MN} \to g_{MN} + h_{MN}$, one obtains

$$S[g+h] = S[g] + \delta S + \delta^2 S + \cdots, \tag{8.2}$$

where first variation is

$$\delta S[g] = -\frac{1}{16\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|}\, h_{MN} \times [-R^{MN} + \tfrac{1}{2} g^{MN}(R + 2\Lambda)], \quad (M, N = 0, 1, 2, \ldots, d+D) \tag{8.3}$$

and the second variation is

$$\delta^2 S[g] = -\frac{1}{16\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|}\, \{(\tfrac{1}{2} g^{MN} h - h^{MN}) \times [\tfrac{1}{2} \Box h_{MN} + \tfrac{1}{2} \nabla_M \nabla_N h - \tfrac{1}{2} g^{PQ} (\nabla_P \nabla_N h_{MQ} + \nabla_M \nabla_P h_{NQ})] + (2h^{MP} h_P^N - h h^{MN}) R_{MN} + [\tfrac{1}{4} h^2 - \tfrac{1}{2} h^{RS} h_{RS}](R + 2\Lambda), \tag{8.4}$$

where h is the trace h_{MN}, $\Box = \nabla^N \nabla_N$ and indices are raised by g^{MN}. If g^{MN} are physical fields satisfying the gravitational field equations

$$\frac{\delta S}{\delta g^{MN}} = 0.$$

In the case of higher dimensional gravity, Killing symmetries (isometries) of the internal space provide gauge group of the theory. So, the gauge transformations are

$$x^M \to x^M + \epsilon^M(x), \tag{8.5}$$

where $\epsilon^M(x)$ are components of the Killing vector in the internal manifold M^D. Here $|\epsilon^M(x)| \ll 1$. As theory is covariant, $S, \delta S$ and $\delta^2 S$ are invariant under gauge transformations given by (8.4). Using Fourier transformations of h and h_{PQ}, one finds that the term

$$(\tfrac{1}{2} g^{MN} h - h^{MN})[\tfrac{1}{2} \nabla_M \nabla_N h - \tfrac{1}{2} g^{PQ} (\nabla_P \nabla_N h_{MQ} + \nabla_M \nabla_P h_{NQ})]$$

yields eigenstates with vanishing eigenvalues. To employ the method of Feynman path integrals, these terms should be removed. For doing so, a suitable gauge-fixing term should be added which is given as

$$S_{GF} = -\frac{1}{32\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|}\quad G_M(g) G_N(g) g^{MN}, \tag{8.6a}$$

where G_M is defined as

$$G_M = g^{PQ} \nabla_P g_{QM} - \tfrac{1}{4} \nabla_M (g^{PQ} g_{PQ}). \tag{8.6b}$$

Terms to be eliminated from $S[g+h]$ appears in the one-loop quantum term $\delta^2 S[g+h]$. So, one-loop term of S_{GF} is also required, which is given as

$$\delta^2 S_{GF} = \frac{1}{32\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|}\{(-\tfrac{1}{2} g^{MN} h + h^{MN}) \times [-\tfrac{1}{2} \nabla_M \nabla_N h + g^{PQ} \nabla_M \nabla_P h_{QN}]\}. \tag{8.7}$$

Adding (8.4) and (8.7), one obtains

$$\begin{aligned} \delta^2 (S[g] + S_{GF}) &= -\frac{1}{16\pi G_{d+D}} \int \mathrm{d}^{d+D}x \sqrt{|g_{(4+D)}|} \\ &\quad \times \{(\tfrac{1}{2} g^{MN} h - h^{MN})[\tfrac{1}{2} \Box h_{MN} - (2h^{MP} h_P^N - h h^{MN})] \\ &\quad \times R_{MN} - [\tfrac{1}{4} h^2 - \tfrac{1}{2} h^{RS} h_{RS}](R + 2\Lambda) + h^{MN} h^{PQ} R_{PMQN}\} \\ &= -\frac{1}{64\pi G_{d+D}} \int \mathrm{d}^{d+D}x \mathrm{d}^{d+D}x' \sqrt{|g_{(4+D)}(x)|} \\ &\quad \times \sqrt{|g_{(4+D)}(x')|}\, h_{MN} \nabla^P \nabla^Q \zeta^{MN} h_{PQ}, \end{aligned} \tag{8.8a}$$

where the operator $\nabla^P \nabla^Q \zeta^{MN}$ is given as

$$\nabla^P \nabla^Q \zeta^{MN} = [\tfrac{1}{2} g^{PQ} g^{MN} - g^{PM} g^{QN}] \Box' - 2(g^{PN} R^{MQ} - g^{PQ} R^{MN}) - R^{MNPQ} \tag{8.8b}$$

with

$$\Box' = \Box - R - 2\Lambda. \tag{8.8c}$$

Here gauge condition has not been used. Instead of it, gauge fixing term has been added. In such a case, Fadeev–Popov method is very convenient. Now it remains to calculate Fadeev–Popov ghost matrix $M^{(G)}$ defined as

$$\int \mathrm{d}^D x' \sqrt{|g|}\, M_Q^{(G)P}(x, x') \epsilon_P(x') = \delta_\epsilon G_Q. \tag{8.9}$$

Under gauge transformations, given by (8.5), change in g_{PQ} can be calculated as follows. It is known that

$$\begin{aligned} g'_{PQ}(x') &= \frac{\partial x^M}{\partial x'^P} \frac{\partial x^N}{\partial x'^Q} g_{MN}(x) \\ &= (\delta_P^M - \epsilon^M_{,P})(\delta_Q^N - \epsilon^N_{,Q})(g_{MN}(x') - \epsilon^R g_{MN,R}) \\ &= g_{PQ}(x') - \epsilon_{P,Q} - \epsilon_{Q,P} - \epsilon^R g_{PQ,R}, \end{aligned} \tag{8.10a}$$

where terms containing higher orders of ϵ are neglected. Equation (8.10a) yields

$$\delta_\epsilon g_{PQ} = \nabla_Q \epsilon_P + \nabla_P \epsilon_Q. \tag{8.10b}$$

The result, given by (8.10b), implies that

$$\delta_\epsilon G_M = \Box \epsilon_M + R^P_M \epsilon_P. \tag{8.11}$$

Now the effective action $\Gamma(g)$ is given by [60]

$$\exp(-\Gamma(g)) = \exp(-S_{cl}[g]) \int D[h_{MN}] \exp(-S_E) \det\left(\frac{\partial G_M}{\partial \epsilon_N}\right), \tag{8.12}$$

where

$$S_E = [\delta^2 (S_{cl} + S_{GF})]_E = \frac{i}{64\pi G_{d+D}} \int \mathrm{d}^{d+D} x \sqrt{|g_{(4+D)}(x)|}\, h_{MN} \nabla^P \nabla^Q \zeta^{MN} h_{PQ} \tag{8.13a}$$

and

$$\det\left(\frac{\partial G_M}{\partial \epsilon_N}\right) = \det(M_M^{(G)N}) = \det(\Box \delta_M^N + R_M^N). \tag{8.13b}$$

Thus

$$\Gamma(g) = S[g] + \tfrac{1}{2} \ln \det \nabla^P \nabla^Q \zeta^{MN} - \ln \det M^{(G)}. \tag{8.14}$$

For evaluation of the integral in (8.12), one can refer to Appendix D.

Determinants in (8.14) can be evaluated on choosing a particular manifold. Chodos and Myers [21] have considered the manifold $M^d \otimes S^D$ (M^d is d-dimensional Minkowski space-time and S^D is D-dimensional sphere) with components of the metric tensor

$$g_{MN} = \eta_{MN} + r^2 \kappa_{MN}, \tag{8.15}$$

where

$$\eta_{MN} = \begin{cases} \eta_{mn} & \text{for } m,n = 0,1,\ldots,(d-1) \\ 0 & \text{for } i,j = d,d+1,\ldots,(d+D-1) \end{cases}$$

and

$$\kappa_{MN} = \begin{cases} 0 & \text{for } m,n = 0,1,\ldots,(d-1) \\ \kappa_{ij} & \text{for } i,j = d,d+1,\ldots,(d+D-1). \end{cases}$$

Here η_{ab} are Minkowski metric components and κ_{ij} are metric components on S^D.

In the manifold $M^d \otimes S^D, M^d$ has no curvature. Since S^D is maximally symmetric, curvature terms are given as

$$R^{MNPQ} = r^{-2}[\kappa^{MN}\kappa^{PQ} - \kappa^{MQ}\kappa^{NP}], \tag{8.16a}$$

$$R^{PQ} = r^{-2}(D-1)\kappa^{PQ}, \tag{8.16b}$$

$$R = r^{-2}D(D-1). \tag{8.16c}$$

If Ξ is the eigenvalue of the operator ζ^{MNPQ}, one can write

$$\int \mathrm{d}^D x' \sqrt{|g(x')|} \times \nabla^P \nabla^Q \zeta^{MN} h_{PQ} = \Xi h^{MN}. \tag{8.17}$$

h_{PQ} are components of a second rank symmetric tensor. So, in the language of differential forms, it can be treated as 1-form (p-form is a covariant anti-symmetric tensor). Now using Hodge's decomposition theorem (refer to Appendix E), h_{PQ} can be decomposed into exact form (longitudinal component), co-exact form (transverse component) and harmonic form. For S^D, Betti number b_0 and b_1 are equal to 1 and other Betti numbers vanish. So, harmonic components of h_{PQ} will vanish as $b_1 = 0$ (it is so because h_{PQ} are treated as 1-forms). On the basis of the above discussion, one can write

$$h_{PQ} = p_{PQ} + \Phi g_{PQ} + \chi \kappa_{PQ}, \tag{8.18a}$$

where p_{PQ}, being exact form, is the longitudinal component and $(\Phi g_{PQ} + \chi \kappa_{PQ})$, being co-exact form, is the transverse component. So

$$\Phi g_{PQ} + \chi \kappa_{PQ} = 0.$$

But Φ and χ are linearly independent, so

$$g^{PQ} p_{PQ} = 0 = \kappa^{PQ} p_{PQ}. \tag{8.18b}$$

Connecting (8.8), (8.17) and (8.18),

$$-\Box' p^{MN} - r^{-2}(D-1)[\kappa^{PN} g^{MQ} + \kappa^{QM} g^{NP}] p_{PQ} + 2r^{-2}\kappa^{MQ}\kappa^{NP} p_{PQ} = \Xi p^{MN}, \tag{8.19a}$$

$$\tfrac{1}{2}(D+d-2)\Box'\Phi+D(D-1)r^{-2}\Phi+\tfrac{1}{2}D\Box'\chi+D(D-1)r^{-2}\chi=\Xi\Phi, \quad (8.19b)$$

$$\Box'\chi-(D-1)(D-4)r^{-2}\chi-(D-1)(D+d-4)r^{-2}\Phi=\Xi\chi. \quad (8.19c)$$

Here $\Box$ is the Laplacian on $M^d \otimes S^D$, so

$$\Box=\Box_m+r^{-2}\Box_S, \quad (8.20)$$

where $\Box_m$ is the Laplacian on M^d and $\Box_S$ is the same on S^D. Obviously, the operator $\Box_m$ has eigenvalues $-\kappa^2$. Moreover, coupled equations (8.19b) and (8.19c) yield

$$(D+d-2)\Box\Phi-2[(D+d-2)\Lambda+\Xi]\Phi + [D(D-1)(D-2)r^{-2}+D\Xi]\chi=0, \quad (8.21a)$$

$$\Box\chi-[2\Lambda+2(D-1)(D-2)r^{-2}+\Xi] - [(D-1)(D+d-4)r^{-2}]\Phi=0. \quad (8.21b)$$

Combining these equations, one obtains

$$(D+d-2)[\Box^2-\{2\Lambda+2(D-1)(D-2)r^{-2}+\Xi\}\Box]\Phi - 2[(D+d-2)\Lambda+\Xi][\Box-\{2\Lambda+2(D-1)(D-2)r^{-2}+\Xi\}]\Phi + (D-1)(D+d-4)r^{-2}[(D-1)(D-2)r^{-2}+\Xi]\Phi=0. \quad (8.22)$$

p^{MN} are scalars on S^D when $M=m$ and $N=n$. These are vector components when $M=i$ (internal index) and $N=n$ and tensor components when $M=i$ and $N=j$. Corresponding to different nature of p^{MN} in three different cases, (8.19a) can be written as

$$[-\Box+2\Lambda+D(D-1)r^{-2}-\Xi]p^{mn}=0, \quad (8.23a)$$

$$[-\Box+2\Lambda+(D-1)^2r^{-2}-\Xi]p^{in}=0, \quad (8.23b)$$

(using $\kappa^{in}=0$) and

$$[-\Box+2\Lambda+(D^2-3D+4)(D-1)r^{-2}-\Xi]p^{ij}=0. \quad (8.23c)$$

Eigenvalue of the operator in (8.23a) is obtained using (4.12) as

$$\Xi=\kappa^2+2\Lambda+[l^2+(D-1)(l+D)]r^{-2}$$

with determinant given by

$$D_1=\prod_{\kappa}\prod_{l=0}^{\infty}[\{l^2+(D-1)(l+D)\}r^{-2}+2\Lambda+\kappa^2]^{(1/2)(d+1)dd_S(l)}, \quad (8.24a)$$

where $d_S(l)$ is given by (4.13). The vector components p^{in} on S^D can be decomposed using Hodge's decomposition theorem as

$$p^{in}=W^{in}+\nabla^i\sigma^n,$$

where W^{in} is the co-exact form and $\nabla^i\sigma^n$ is an exact form (σ^n are scalars). The eigenvalue of the operator in (8.23b) corresponding to W^{in} is

$$\Xi = \kappa^2 + 2\Lambda + [l^2 + (D-1)(l+D-1) - 1]r^{-2}$$

with determinant

$$D_2 = \prod_{\kappa}\prod_{l=1}^{\infty}[\{l^2 + (D-1)(l+D-2)\}r^{-2} + 2\Lambda + \kappa^2]^{dd_v(l)}, \tag{8.24b}$$

where $d_v(l)$ is given by (4.15). Moreover,

$$\begin{aligned}\Box(\nabla^i\sigma^n) &= \nabla^i(\Box\sigma^n) + R^{il}\nabla_l\sigma^n \\ &= [-\kappa^2 - l(l+D-1)r^{-2}]\nabla^i\sigma^n + (D-1)r^{-2}\nabla^i\sigma^n \\ &= [-\kappa^2 - \{l^2 + l - 1)(D-1)r^{-2}\}]\nabla^i\sigma^n.\end{aligned}$$

Now Ξ for $\nabla^i\sigma^n$ is given as

$$\Xi = \kappa^2 + 2\Lambda + [l^2 + (D-1)(l+D-2)]r^{-2}.$$

So, the determinant is

$$D_3 = \prod_{\kappa}\prod_{l=1}^{\infty}[\{l^2 + (D-1)(l+D-2)\}r^{-2} + 2\Lambda + \kappa^2]^{dd_v(l)}. \tag{8.24c}$$

Three cases are possible for tensor components p^{ij}. Using (4.16) in (8.23c) for divergenceless tensor,

$$D_4 = \prod_{\kappa}\prod_{l=2}^{\infty}[\{l^2 + (D-1)(l+D-2)\}r^{-2} + 2\Lambda + \kappa^2]^{d_T(l)}. \tag{8.24d}$$

The other two cases yield determinants

$$D_5 = \prod_{\kappa}\prod_{l=2}^{\infty}[\{l^2 + (D-1)(l+D-3)\}r^{-2} + 2\Lambda + \kappa^2]^{d_v(l)}, \tag{8.24e}$$

$$D_6 = \prod_{\kappa}\prod_{l=2}^{\infty}[\{l^2 + (D-1)(l+D-4)\}r^{-2} + 2\Lambda + \kappa^2]^{d_S(l)}. \tag{8.24f}$$

The determinant of eigenvalues for the operator in (8.22) is given as

$$\begin{aligned}D_7 = \prod_{\kappa}\prod_{l=1}^{\infty}\{(D+d-2)[\{l^2 + (D-1)(l+D-2)\}r^{-2} + 2\Lambda + \kappa^2] \\ + 2(D-1)^2(D-2)(d-2)r^{-4}\}^{d_S(l)}.\end{aligned} \tag{8.24g}$$

The ghost determinant $\det M^{(G)}$ assumes a similar form with eigenvalue equation

$$(\Box\delta^N_M + R^N_M)\epsilon_N = \Xi_G\epsilon_N,$$

which can be rewritten as

$$(\Box\delta_M^N + (D-1)r^{-2}\kappa_M^N)\epsilon_N = \Xi_G\epsilon_N. \tag{8.25}$$

When $N = n, \epsilon_n$ is scalar on S^D. So, the determinant of eigenvalue

$$D_1^{(G)} = \prod_{\kappa}\prod_{l=0}^{\infty}[l(l+D-1)r^{-2} + \kappa^2]^{dd_S(l)}. \tag{8.26a}$$

When $N = i, \epsilon_i$ are components of a vector on S^D, which can be decomposed into exact and co-exact forms using Hodge's decomposition theorem as above. Now the determinant of eigenvalues are

$$D_2^{(G)} = \prod_{\kappa}\prod_{l=1}^{\infty}[\{l^2 + l(D-1) + (D-2)\}r^{-2} + \kappa^2]^{d_S(l)} \tag{8.26b}$$

and

$$D_3^{(G)} = \prod_{\kappa}\prod_{l=1}^{\infty}[\{l^2 + l(D-1)\}r^{-2} + \kappa^2]^{d_v(l)}. \tag{8.26c}$$

One can notice that all these determinants, given by (8.24) and (8.26), have a general form

$$D = \prod_{\kappa}\prod_{l=l_0}^{\infty}[\{l^2 + l(D-1) + (D-2) + c(x,D)\}r^{-2} + \kappa^2]^{d(l)}. \tag{8.27}$$

As a result

$$\ln D = -\frac{\mathrm{d}\zeta}{\mathrm{d}s}\Big|_{s=0}. \tag{8.28}$$

Now $\zeta(s)$ is evaluated in the region where s is large and the result is analytically continued to $s = 0$. For evaluation of $\zeta(s)$, one can refer to Candelas and Weinberg [26] as well as Chodos and Myers [21].

As explained above, evaluation of these determinants yields one-loop quantum correction V_Q given as

$$V_Q = (2\pi r)^4\left[-\frac{15}{2}\left\{\sum_{n=1}^{\infty}\mathrm{e}^{(-2\pi nx)}\left(\frac{3}{2\pi^2n^5} + \frac{3x}{\pi n^4} + \frac{2x^2}{n^3}\right) - \frac{1}{\pi^2}\zeta(5)\right\} + 4\pi^3x^3\right]. \tag{8.29}$$

In 1990, Buchbinder *et al* have calculated 2-loop effective action on the background geometry $M^d \otimes T^D$ [61].

9. One-loop quantum correction and induced gravity

In 1983, using 5-dimensional space-time with topology $M^4 \otimes S^1$ and line-element, given by (3.2), for $\phi(x) = 1$, Toms [62] obtained induced gravitational action as well as

Maxwell's action through one-loop quantum correction to scalar and spinor fields. The action for the scalar field $\hat{\phi}$ is taken as

$$S_{\hat{\phi}} = \frac{1}{2}\int \mathrm{d}^5x\sqrt{|\hat{g}|}[g^{\hat{\mu}\hat{\nu}}\partial_{\hat{\mu}}\hat{\phi}\partial_{\hat{\nu}}\hat{\phi} - m^2\hat{\phi}^2 - \xi\hat{R}_5\hat{\phi}^2], \tag{9.1}$$

where ξ is the non-minimal coupling constant, S^1 is a non-simply connected manifold. So $\hat{\phi}(x,y)$ can be twisted or untwisted field. Untwisted field is periodic in y and twisted field is anti-periodic in y [63, 64]. $\hat{\phi}(x,y)$ can be decomposed as

$$\hat{\phi}(x,y) = \sum_{n=-\infty}^{\infty}\hat{\phi}_{(n)}(x)\exp[2\pi(n+\alpha)y/L], \tag{9.2a}$$

where

$$\alpha = \begin{cases} 0 & \text{for untwisted field} \\ \frac{1}{2} & \text{for twisted field.} \end{cases} \tag{9.2b}$$

Connecting (9.1) and (9.2) and integrating over y

$$\begin{aligned} S_{\hat{\phi}} = &\frac{1}{2}\sum_{-\infty}^{\infty}\int \mathrm{d}^4x\sqrt{|g|}\,[g^{\mu\nu}(D_\mu^{(n)}\hat{\phi}_{(n)})^*(D_\nu^{(n)}\hat{\phi}_{(n)}) \\ &- M_{(n)}^2\hat{\phi}_{(n)}^*\hat{\phi}_{(n)} - \xi R\hat{\phi}_{(n)}^*\hat{\phi}_{(n)} + \tfrac{1}{4}\xi\kappa^2F_{\mu\nu}F^{\mu\nu}\hat{\phi}_{(n)}^*\hat{\phi}_{(n)}], \end{aligned} \tag{9.3a}$$

where

$$\begin{aligned} D_\nu^{(n)}\hat{\phi}_{(n)} &= (\partial_\mu - iq_{(n)}A_\mu)\phi_{(n)}, \\ M_{(n)}^2 &= m^2 + 4\pi^2(n+\alpha)^2L^{-2}, \\ q_{(n)} &= 2\pi\kappa(n+\alpha)L^{-1}. \end{aligned} \tag{9.3b}$$

The one-loop effective action can be written, in this case, as

$$\Gamma^{(1)} = \frac{i}{2}\sum_{-\infty}^{\infty}\ln\det\triangle_{(n)}, \tag{9.4a}$$

where the operator

$$\triangle_{(n)} = g^{\mu\nu}D_\mu^{(n)}D_\nu^{(n)} + M_{(n)}^2 + \xi R - \tfrac{1}{4}\xi\kappa^2F_{\mu\nu}F^{\mu\nu}. \tag{9.4b}$$

Toms has computed $\Gamma^{(1)}$, using the heat-kernel method where kernel $K_{(n)}(s,x,x')$ is defined for the operator $\triangle_{(n)}$ [56]. As a result

$$\Gamma^{(1)} = -\frac{i}{2}\sum_{-\infty}^{\infty}\int \mathrm{d}^4x\sqrt{|g|}\int_0^\infty\frac{\mathrm{d}s}{s}\,\mathrm{tr}\,K_{(n)}(s,x,x'), \tag{9.5a}$$

where tr stands for the trace. The Schwinger DeWitt expansion of $\mathrm{tr}\,K_n(s,x,x')$ yields

$$\mathrm{tr}\,K_n(s,x,x') = i\mu^{4-N}(4\pi is)^{-N/2}\exp(-iM_n^2s)\sum_{p=0}^{\infty}(is)^pa_p(x), \tag{9.5b}$$

where μ is a parameter of mass dimension and N is a parameter for dimensional regularization such that one gets the actual 4-dimensional theory on taking the limit $N \to 4$. Connecting (9.5a) and (9.5b) one can find that the first three terms in the series yields divergence in the integral

$$\int_0^\infty \frac{\mathrm{d}s}{s} i\mu^{4-N}(4\pi is)^{-N/2} \exp(-iM_n^2 s)$$
$$\times [a_0(x,x') + (is)a_1(x,x') + (is)^2 a_2(x,x') + \cdots]. \tag{9.5c}$$

All other terms are convergent for a 4-dimensional theory. So, the coefficients a_0, a_1 and a_2 are crucial. On taking the limit $x \to x'$, these coefficients are given as

$$a_0(x) = 1,$$
$$a_1(x) = (\tfrac{1}{6} - \xi)R + \tfrac{1}{4}\kappa^2 \xi F^{\mu\nu}F_{\mu\nu},$$
$$a_2(x) = -\tfrac{1}{12} q_n^2 F^{\mu\nu}F_{\mu\nu} + \cdots. \tag{9.5d}$$

From (9.5), one obtains for N_0^+ untwisted fields

$$\Gamma^{(1)} = \frac{N_0^+}{4\pi^2}\int \mathrm{d}^4x\sqrt{|g|}$$
$$\times \left[\frac{3\zeta(5)}{L^4} + \frac{\zeta(3)}{2L^2}\left(\frac{1}{6} - \xi\right)R + \left(\frac{1}{3} + \xi\right)\frac{\zeta(3)}{4L^2}\kappa^2 F^{\mu\nu}F_{\mu\nu} + \cdots\right]. \tag{9.6}$$

Other terms are not calculated here, because the purpose is to obtain induced gravity and induced Macwell's term.

For the N_0^- twisted fields ($\alpha = \frac{1}{2}$),

$$\Gamma^{(1)} = -\frac{N_0^-}{32\pi^2}\int \mathrm{d}^4x\sqrt{|g|}$$
$$\times \left[\frac{45\zeta(5)}{2L^4} + \frac{3\zeta(3)}{L^2}\left(\frac{1}{6} - \xi\right)R + \left(\frac{1}{3} + \xi\right)\frac{3\zeta(3)}{4L^2}\kappa^2 F^{\mu\nu}F_{\mu\nu} + \cdots\right]. \tag{9.7}$$

For calculation of these terms, one can refer to Appendix F.

Including $N_{1/2}^+$ untwisted Dirac spinors and $N_{1/2}^-$ twisted spinors also in the theory, the total one-loop effective action

$$\Gamma^{(1)} = \frac{1}{8\pi^2}\int \mathrm{d}^4x\sqrt{|g|}\left[-\frac{45\zeta(5)}{8L^4}\left\{(N_0^- - 4N_{1/2}^-) - \frac{16}{15}(N_0^+ - 4N_{1/2}^+)\right\}\right.$$
$$+\frac{\zeta(3)}{3L^2}\left\{\left(N_{1/2}^+ - \frac{3}{4}N_{1/2}^-\right)\right\} + 3\left(\frac{1}{6} - \xi\right)\left(N_0^+ - \frac{3}{4}N_0^-\right)R - \frac{5\zeta(3)}{24L^2}\kappa^2$$
$$\left.\times\left\{\left(N_{1/2}^+ - \frac{3}{4}N_{1/2}^-\right) - \frac{6}{5}\left(\frac{1}{3} + \xi\right)\left(N_0^+ - \frac{3}{4}N_0^-\right)\right\}F^{\mu\nu}F_{\mu\nu} + \cdots\right] \tag{9.8}$$

This result shows that the induced cosmological constant can vanish provided that $N_0^+ = 4N_{1/2}^+$ and $N_0^- = 4N_{1/2}^-$.

In 1984, Candelas and Weinberg [26] have calculated induced cosmological constant, induced gravity and fine-structure constant using the background geometry of $(4+N)$-dimensional space-time with topology $M^4 \otimes B^N$. Moreover they have calculated one-loop effective potential and discussed stability problem.

Yoshimura [51] has computed the one-loop quantum effective action at finite temperature in the background geometry of higher-dimensional Robertson–Walker type model with topology $R^1 \otimes S^{d_1} \otimes S^{d_2}$. The line-element is taken as

$$dS^2 = d\tau^2 + a_1^2(\tau)\gamma_{ab}^{(1)}dx^a dx^b + a_2^2(\tau)\gamma_{mn}^{(2)}dy^m dy^n, \tag{9.9}$$

where $\tau = it$ (t is the cosmic time), the scale factor a_i is periodic i.e. $a_i(\tau+\beta) = a_i(\tau)$ with $\beta = T^{-1}$. Here scale factors are slowly varying to maintain thermal equillibrium up to a good approximation.

The effective 4-dimensional gravitational constant G_{eff} is obtained here as

$$G_{\text{eff}}^{-1} \simeq G_4^{-1} - \frac{5}{6}\pi^{-(d_2+2)/2}d_2(d_2+1)\Gamma\left(\frac{d_2+2}{2}\right)T^{d_2+2}, \tag{9.10}$$

where G_4 is the Newtonian gravitational constant. It is interesting to note that $G_{\text{eff}} < 0$ above the critical temperature given as

$$T_{\text{cr}} \simeq \left[\frac{6}{5}\pi^{+(d_2+2)/2}\left\{d_2(d_2+1)\Gamma\left(\frac{d_2+2}{2}\right)G_4\right\}^{1/(d_2+2)}\right]. \tag{9.11}$$

Physically, it means that gravity becomes repulsive when $T > T_{\text{cr}}$. In natural units $(\hbar = c = 1)G_4 \simeq M_{\text{P}}^{-2}$($M_{\text{P}}$ is the Planck mass), so T_{cr} will be much below Planck energy if d_2 is large.

Assuming infinitely many cycles of expansion and contraction of the universe, Yoshimura has discussed that temperature might have raised above the critical temperature during contraction and due to anti-gravity effect of $G_{\text{eff}} < 0$, the present expansion might have started. He has also suggested that a huge amount of entropy might have produced due to many cycles of the universe possibly solving horizon and flatness problems of the standard cosmology.

In 1985, Gleiser and Taylor [38] have obtained time-dependence of coupling constants in the six-dimensional cosmological models with topology $M^4 \otimes S^2$. Thus, first time dynamical model was used for these kinds of investigations.

In [41], a gauge-dependent model with topology $M^4 \otimes T^D$ has been considered and solutions of higher-dimensional Einstein's equations have been obtained with energy momentum tensor of scalar as well as Maxwell's fields. Using the background geometry of these solutions, time-dependence of gravitational constant, cosmological constant and fine-structure constants are obtained through dimensional reduction and one-loop quantum correction to scalar fields. Some mathematical details of this approach is given as follows.

The line-element for the underlying model is given by (6.5) with solution of Einstein's equations, given by (6.15). Using the horizontal lift basis, discussed in § 3 in

the case of 5-dimensional theory, the action for pure gravity can be written as

$$S_g = \frac{1}{16\pi G_{4+D}} \int \mathrm{d}^4x\, \mathrm{d}^D y \sqrt{|g|}(R_4 - \tfrac{1}{4}\kappa^2 \delta_{mn} F^m_{\mu\nu} F^{n\mu\nu}), \tag{9.12}$$

where Ricci scalar vanishes on T^D.

Components of the metric tensor g_{MN} on the manifold $M^4 \otimes T^D$ can be written as

$$g_{MN} = \begin{pmatrix} g_{\mu\nu} & 0 \\ 0 & -R_D^2(t)\delta_{mn} \end{pmatrix}, \tag{9.13}$$

where $g_{\mu\nu} \equiv \mathrm{diag}\,(1, -R_3^2, -R_3^2, -R_3^2)$ on M^4 and g_{MN} are metric tensor components on T^D. Now conformal transformation yields

$$g_{MN} \to R_D^2(t)\, \bar{g}_{MN} = \begin{pmatrix} \bar{g}_{\mu\nu} & 0 \\ 0 & -\delta_{mn} \end{pmatrix}, \tag{9.14}$$

where $\bar{g}_{\mu\nu} \equiv \mathrm{diag}\,(1, -R_D^{-2}R_3^2, -R_D^{-2}R_3^2, -R_D^{-2}R_3^2)$.

On using conformal transformations, given by (4.2), in (4.1) double divergence terms for b will appear. Using Gauss's divergence theorem on such terms, one can realize that these terms do not contribute to the theory. As a result, one obtains 4-dimensional action for gravity alongwith Maxwell's term as

$$S_g^{(4)} = -\frac{(2\pi)^2 \rho_1 \rho_2 \cdots \rho_D}{16\pi G_{4+D}} \int \mathrm{d}^4x \sqrt{|\bar{g}_4|} R_D^{D-2} \times \left[\tilde{R}_4 - \frac{D}{4}\kappa^2 \kappa^D \tilde{F}_{\mu\nu}\tilde{F}^{\mu\nu} - (D+2)(D+3)\left(\frac{\dot{R}_D}{R_D}\right)^2\right], \tag{9.15}$$

on integrating over y.

Conformal transformations, given by (9.14), were used for mathematical convenience. So, to come back to the original system another conformal transformation is done to undo the earlier one which are given as

$$\tilde{g}_{\mu\nu} \to R_D^{-2}(t) g_{\mu\nu}. \tag{9.16}$$

Using these transformations, $S_g^{(4)}$ (given by (9.15)) is written as

$$S_g^{(4)} = -\frac{1}{16\pi G_4} \int \mathrm{d}^4x (R_3)^3 (R_D)^D \left[R_4 - \frac{D}{4}\kappa^2 F_{\mu\nu}F^{\mu\nu} - D(D-1)\left(\frac{\dot{R}_D}{R_D}\right)^2\right], \tag{9.17}$$

where $G_4 \equiv G_{4+D}/(2\pi)^2 \rho_1 \rho_2 \cdots \rho_D$.

The action for scalar field $\hat{\phi}$ is given by (6.6). Further it is decomposed into homogeneous part $\hat{\phi}_0$ and inhomogeneous part $\tilde{\phi}$ as given by (6.11). For convenience, one can write

$$\hat{\phi}_0 = [(2\pi)^2 \rho_1 \rho_2 \cdots \rho_D (R_D)^D]^{(-1/2)} \tilde{\hat{\phi}}_0 \tag{9.18a}$$

and

$$\tilde{\bar{\phi}}=[(2\pi)^2\rho_1\rho_2\cdots\rho_D(R_D)^D]^{(-1/2)}\sum_{n_1n_2\cdots n_D=-\infty}^{\infty}\tilde{\hat{\phi}}_{(n)}\exp\left[i\sum_{j=1}^{D}\{2\pi(n_j+\alpha)y_j/\rho_j\}\right], \tag{9.18b}$$

where α is given by (9.2b) and $\tilde{\hat{\phi}}_{(n)}\equiv\tilde{\hat{\phi}}_{n_1n_2\cdots n_D}$. Connecting (6.6) and (9.18) and integrating over y, it is obtained that

$$S(\phi)=S^{(4)}(\tilde{\hat{\phi}}_0)+S^{(4)}(\tilde{\hat{\phi}}_n), \tag{9.19a}$$

where

$$S^{(4)}(\tilde{\hat{\phi}}_0)=\frac{1}{2}\int \mathrm{d}^4x(R_3)^3\tilde{\hat{\phi}}_0\left[-\Box_4+m^2+\frac{3D}{2}\left(\frac{\dot{R}_3}{R_3}\right)\left(\frac{\dot{R}_D}{R_D}\right)+D\left(\frac{\dot{R}_D}{R_D}\right)^2+\frac{D}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{R}_D}{R_D}\right)\right]\tilde{\hat{\phi}}_0 \tag{9.19b}$$

and

$$S^{(4)}(\tilde{\hat{\phi}}_{(n)})=\frac{1}{2}\sum_{n_1n_2\cdots n_D=-\infty}^{\infty}\int \mathrm{d}^4x(R_3)^3\times[g^{\mu\nu}(D_\mu\tilde{\hat{\phi}}_{(N)})^*(D_\mu\tilde{\hat{\phi}}_{(N)})-M^2_{(n)}\tilde{\hat{\phi}}^*_{(N)}\tilde{\hat{\phi}}_{(N)}], \tag{9.19c}$$

where

$$M^2_{(n)}=\left(\frac{2\pi}{R_D}\right)^2\sum_{j=1}^{D}\{2\pi(n_j+\alpha)/\rho_j\}^2+m^2+\frac{3D}{2}\left(\frac{\dot{R}_3}{R_3}\right)\left(\frac{\dot{R}_D}{R_D}\right)+D\left(\frac{\dot{R}_D}{R_D}\right)^2+\frac{D}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{R}_D}{R_D}\right) \tag{9.19d}$$

and

$$D_\mu\tilde{\hat{\phi}}_{(n)}=\nabla_\mu\tilde{\hat{\phi}}_{(n)}+iq_{(n)}A_\mu\tilde{\hat{\phi}}_{(n)} \tag{9.19e}$$

with

$$q_{(n)}=\kappa M\sum_{j=1}^{D}n_j,\quad M=\frac{2\pi}{\rho_H}$$

and

$$\rho_H=\sum_{j=1}^{D}n_j\Big/\sum_{j=1}^{D}(n_j/\rho_j).$$

Recognizing κM as e, one finds $q_{(n)}$ as charge of scalar field $\tilde{\hat{\phi}}_{(n)}$ being integral multiple of e.

In curved spaces, operator regularization method is very convenient, because it leads to finite results [65,66]. Using this method, one-loop quantum correction to $\tilde{\hat{\phi}}_{(n)}$ is computed and the series is summed up. Up to adiabatic order 4, the effective action with one-loop quantum correction is given as

$$\Gamma = S(\tilde{\hat{\phi}}) + \sum_{n_1 n_2 \cdots n_D = -\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}s}\left[\left(\frac{\mu^2}{M^2_{(n)}}\right)^s \int \mathrm{d}^4x (R_3)^2 \right.$$
$$\times \left\{ \frac{M^4_{(n)}}{(s-1)(s-2)} + \frac{M^2_{(n)}}{6(s-1)} R_4 + \left(\frac{1}{30} \Box_{(4)} R_{(4)} + \frac{1}{180} R^{\mu\nu\alpha\beta}_{(4)} R_{(4)\mu\nu\alpha\beta} \right.\right.$$
$$\left.\left. \left. - \frac{1}{180} R^{\mu\nu}_{(4)} R_{(4)\mu\nu} + \frac{1}{72} R^2_{(4)} + \frac{D}{12} F_{\mu\nu} F^{\mu\nu} \right) \right\} \Bigg] \Bigg|_{s=0}$$

which reduces to

$$\Gamma = S(\tilde{\hat{\phi}}) + \int \mathrm{d}^4x (R_3)^2 \left[\ln(\mu^2/m^2) \left\{ \frac{m^4}{2} - \frac{m^2}{6} R_{(4)} + \frac{1}{30} \Box_{(4)} R_{(4)} \right.\right.$$
$$\left. + \frac{1}{180} R^{\mu\nu\alpha\beta}_{(4)} R_{(4)\mu\nu\alpha\beta} - \frac{1}{180} R^{\mu\nu}_{(4)} R_{(4)\mu\nu} + \frac{1}{72} R^2_{(4)} + \frac{D}{12} \kappa^2 F_{\mu\nu} F^{\mu\nu} \right\}$$
$$\left. + \left\{ \frac{3m^4}{4} - \frac{m^2}{6} R_{(4)} \right\} \right]. \tag{9.20}$$

For summation of series, one can refer to Appendix F.

Equations (9.17) and (9.20) yield the effective Maxwell's term as

$$\int \mathrm{d}^4x (R_3)^2 \left[\frac{D\kappa^2 (R_D)^D}{64\pi G_4} + \frac{D\kappa^2}{12} \ln(\mu^2/m^2) \right] F_{\mu\nu} F^{\mu\nu}.$$

Using normalization of electromagnetic fields, one obtains

$$\left[\frac{D\kappa^2 (R_D)^D}{64\pi G_4} + \frac{D\kappa^2}{12} \ln(\mu^2/m^2) \right] = 1, \tag{9.21}$$

as the standard Maxwell's term in the action is

$$\frac{1}{4} \int \mathrm{d}^4x \sqrt{|g|} F_{\mu\nu} F^{\mu\nu}.$$

Since $\kappa = e/M = e\rho_H/2\pi$, (9.21) implies that

$$\frac{e^2}{4\pi} = \left[\frac{D\rho_H^2}{\pi} \left\{ \frac{(R_D)^D}{16\pi G_4} + \frac{1}{3} \ln(\mu^2/m^2) \right\} \right]^{-1}. \tag{9.22}$$

Equations (9.17) and (9.20) yield the effective action for 4-dimensional gravity as

$$S_g^{(4)} = -\int \mathrm{d}^4x (R_3)^3 \left[\left\{ \frac{(R_D)^D}{16\pi G_4} + \frac{m^2}{6} + \frac{m^2}{6} \ln(\mu^2/m^2) \right\} R_4 \right.$$
$$\left. - \left\{ D(D-1) \left(\frac{\dot{R}_D}{R_D} \right)^2 + \frac{3m^4}{4} + \frac{m^4}{2} \ln(\mu^2/m^2) \right\} \right], \tag{9.23}$$

which implies that

$$\frac{1}{16\pi G_{\text{eff}}} = \frac{(R_D)^D}{16\pi G_4} + \frac{m^2}{6} + \frac{m^2}{6} \ln(\mu^2/m^2) \tag{9.24a}$$

and

$$\frac{\Lambda_{\text{eff}}}{8\pi G_{\text{eff}}} = \left\{ D(D-1) \left(\frac{\dot{R}_D}{R_D} \right)^2 + \frac{3m^4}{4} + \frac{m^4}{2} \ln(\mu^2/m^2) \right\}. \tag{9.24b}$$

Renormalization at $\mu^2 = m^2$ leads to the following results from (9.22) and (9.24) given as

$$\frac{e^2}{4\pi} = \frac{16\pi^2 G_4}{D\rho_H^2 (R_D)^D}, \tag{9.25a}$$

$$\frac{1}{16\pi G_{\text{eff}}} = \frac{(R_D)^D}{16\pi G_4} + \frac{m^2}{6} \tag{9.25b}$$

and

$$\frac{\Lambda_{\text{eff}}}{8\pi G_{\text{eff}}} = D(D-1) \left(\frac{\dot{R}_D}{R_D} \right)^2 + \frac{3m^4}{4}. \tag{9.25c}$$

Using solutions of Einstein's field equations, given by (6.15), one obtains time-dependence of $e^2/4\pi, G_{\text{eff}}$ and Λ_{eff}. It is interesting to note that when $t \to \infty$, time-dependence of these terms ceases to be effective. In [42], time-dependence of fundamental constants has been discussed in singularity-free 5-dimensional Kaluza–Klein cosmological model. In this paper, heat-kernel method has been used for calculation of one-loop correction. In [67, 68], $(1+1)$-dimensional as well as $(1+2)$-dimensional cosmological models have been obtained using these methods.

Appendix A

Dimensional reduction of Dirac spinors

In a space-time with topology $M^{d-D} \otimes K^D$ where K^D is a compact manifold with space-like coordinates y^a and M^{d-D} is the $(d-D)$-dimensional Minkowskian space-time with coordinates x^μ. Dirac lagrangian for a spinor $\psi(x, y)$ is given as

$$L_\psi = \tfrac{1}{2} \bar{\psi} [i(\gamma^\mu \nabla_\mu + \gamma^a \nabla_a) + m_d] \psi + \text{h.c.} \tag{A1}$$

ψ can be expanded as [21, 22, 66]

$$\psi(x,y) = \sum_i \psi_M(x)\psi_{iA}(y), \tag{A2}$$

where $A = 1, \ldots, \bar{A}$ and $M = 1, \ldots, \bar{M}$ such that $\bar{A}\bar{M} = 2^{[d/2]}$. The appropriate choice for the d-dimensional Dirac matrices $\gamma^{\bar{\mu}} \equiv (\gamma^\mu, \gamma^a)$ are given by [69]

$$\gamma^\mu = \Gamma^\mu \otimes S_{\bar{A}}, \quad \gamma^a = \Gamma^a \otimes S_{\bar{M}} \tag{A3}$$

with $S_{\bar{A}}S_{\bar{A}} = I_{\bar{A}}$, $S_{\bar{M}}S_{\bar{M}} = I_{\bar{M}}$, $S_{\bar{A}}^\dagger = S_{\bar{A}}$, $S_{\bar{M}}^\dagger = S_{\bar{M}}$ and $\{S_{\bar{A}}, \Gamma^a\} = 0$, $[S_{\bar{M}}, \Gamma^\mu] = 0$ or $[S_{\bar{A}}, \Gamma^a] = 0$, $\{S_{\bar{M}}, \Gamma^\mu\} = 0$. If $(d-D)$ is even, one can choose $\bar{A} = 2^{[d-D]/2}$, $\bar{M} = 2^{[D]/2}$, $S_{\bar{A}} = I_{\bar{A}}$, $S_{\bar{M}} = \eta\Gamma^0\Gamma^1\cdots\Gamma^{d-D-1}$. If $(d-D)$ is odd, Γ^μ are $2^{[d-D-1]/2}$ in number and $S_{\bar{M}}$ is the last Γ^μ i.e. Γ^{d-D-1} which implies that

$$\begin{aligned} S_{\bar{M}} &= \eta\Gamma^0\Gamma^1\cdots\Gamma^{d-D-1} = \eta\Gamma^0\Gamma^1\cdots\Gamma^{d-D-2}S_{\bar{M}} \\ &= \eta\Gamma^0\Gamma^1\cdots\Gamma^{d-D-2}\eta\Gamma^0\Gamma^1\cdots\Gamma^{d-D-2} = I_{\bar{M}}. \end{aligned}$$

Thus $S_{\bar{M}}$ commutes with all Γ^μ in case $(d-D)$ is odd. If D is odd, $S_{\bar{A}} = I_{\bar{A}}$ with $\bar{A} = 2^{[D-1]/2}$.

In (A2), $\psi_{iA}(y)$ satisfy the equation

$$\Gamma^a \nabla_a \psi_{iA}(y) = M\psi_{iA}(y), \tag{A4}$$

where M is the mass matrix. Connecting (A1), (A2) and (A4)

$$L_\psi = \tfrac{1}{2}\bar{\psi}_M[i\gamma^\mu\nabla_\mu + (m_d + iMS_{\bar{M}})]\psi_M + \text{h.c.} \tag{A5}$$

It is difficult to handle the mass term obtained here. So, for convenience, a transformation is used which is given as

$$\psi_M \to \psi'_M = \exp[i\alpha S_{\bar{M}}]\psi_M. \tag{A6}$$

Also,

$$\begin{aligned} \exp[i\alpha S_{\bar{M}}] &= 1 + i\alpha S_{\bar{M}} + \frac{(i\alpha S_{\bar{M}})^2}{2!} + \frac{(i\alpha S_{\bar{M}})^3}{3!} + \cdots \\ &= \cos\alpha + iS_{\bar{M}}\sin\alpha. \end{aligned} \tag{A7}$$

From (A6), one obtains

$$\bar{\psi}_M = \psi_M^\dagger\Gamma^0 = \bar{\psi}'_M(\cos\alpha - iS_{\bar{M}}\sin\alpha). \tag{A8}$$

Now

$$\begin{aligned} &\bar{\psi}_M[i\Gamma^\mu\nabla_\mu + (m_d + iMS_{\bar{M}})]\psi_M \\ &\quad = \bar{\psi}'_M(\cos\alpha - iS_{\bar{M}}\sin\alpha)(\cos\alpha + iS_{\bar{M}}\sin\alpha)\Gamma^\mu\nabla_\mu\psi'_M \\ &\qquad + \bar{\psi}'_M[(m_d\cos 2\alpha + M\sin 2\alpha) + iS_M(M\cos 2\alpha - m_d\sin 2\alpha)]\psi'_M \\ &\quad = \bar{\psi}'_M[\Gamma^\mu\nabla_\mu + ((m_d)^2 + M^2)^{1/2}]\psi'_M, \end{aligned} \tag{A9}$$

as $\tan 2\alpha = M/m_d$.

Appendix B

Complex projective space

If $z^k, k = 1, 2, \ldots, N+1$ are complex numbers and λ is another complex number such that

$$(z^1, z^2, \ldots, z^{N+1}) = (\lambda z^1, \lambda z^2, \ldots, \lambda z^{N+1}),$$

the resulting space is the complex projective space CP^N. It is a simply connected compact manifold. In case $N = 1$,

$$(z^1, z^2) = (\lambda z^1, \lambda z^2),$$

which gives CP^1 or S^2. We have

$$(z^1, z^2, z^3) = (\lambda z^1, \lambda z^2, \lambda z^3),$$

in case $N = 2$, which gives CP^2. The coordinates on CP^2 are z^1/z^3 and z^2/z^3. Thus $(z^1/z^3,\ z^1/z^3)$ cover all points on CP^2 barring points for which $z^3 = 0$, such regions are homeomorphic to R^4 (deformable to R^4). Points, for which, $z^3 = 0$ may be regarded as points at infinity. Such points belong to CP^1. If

$$|z^1|^2 + |z^2|^2 + |z^3|^2 = \frac{6}{\lambda},$$

the line-element on CP^2 is given as

$$\mathrm{d}S^2 = \left[1 + \frac{\lambda}{6}\left(|\xi^1|^2 + |\xi^2|^2\right)\right]^{-2}\left[|\mathrm{d}\xi^1|^2 + |\mathrm{d}\xi^1|^2 + \frac{\lambda}{6}|\mathrm{d}(\xi^1\xi^{-2})|^2 - \frac{\lambda}{6}\,\mathrm{d}(|\xi^1|^2)\mathrm{d}(|\xi^2|^2)\right],$$

where $\xi^1 = z^1/z^3$ and $\xi^2 = z^2/z^3$.

Appendix C

Adiabatic regularization

Adiabatic regularization method was developed by Parker *et al* for regularization in Robertson–Walker space-times,

$$\mathrm{d}S^2 = \mathrm{d}t^2 - a_1^2(t)\mathrm{d}x^2 - a_2^2(t)\mathrm{d}y^2 - a_3^2(t)\mathrm{d}z^2. \tag{C1}$$

The matter field is provided by neutral scalar field ϕ with the lagrangian

$$L = \tfrac{1}{2}\sqrt{|g|}\{g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \tfrac{1}{6}R\phi^2 - m^2\phi^2\}, \tag{C2}$$

which yields the field equation for ϕ in the background geometry, given by (C1), as

$$\ddot{\phi} + (\dot{V}/V)\dot{\phi} - \frac{1}{a_1^2}\frac{\partial^2}{\partial x^2}\phi - \frac{1}{a_2^2}\frac{\partial^2}{\partial y^2}\phi - \frac{1}{a_3^2}\frac{\partial^2}{\partial z^2}\phi + \left(\frac{1}{6}R + m^2\right)\phi = 0, \tag{C3}$$

where $V = \sqrt{|g|} = a_1 a_2 a_3$ and dot denotes partial differentiation with respect to time t. The equal-time canonical commutation relations are

$$[\phi(\mathbf{x}, t), \phi(\mathbf{x}', t)] = [\pi(\mathbf{x}, t), \pi(\mathbf{x}', t)] = 0, \tag{C4a}$$

$$[\phi(\mathbf{x}, t), \pi(\mathbf{x}', t)] = i\delta(\mathbf{x} - \mathbf{x}'), \tag{C4b}$$

where $\pi = \partial L / \partial \dot{\phi} = V\dot{\phi}$.

Components of the energy-momentum tensor are given by

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \tfrac{1}{2} g_{\mu\nu} g^{\lambda\sigma} \partial_\mu \phi \partial_\nu \phi + \tfrac{1}{2} g_{\mu\nu} m^2 \phi^2 - \tfrac{1}{6} \nabla_\mu \partial_\nu (\phi^2) + \tfrac{1}{6} g_{\mu\nu} g^{\lambda\sigma} \nabla_\lambda \partial_\sigma (\phi^2) - \tfrac{1}{6} (\phi^2)(R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R). \tag{C5}$$

Equation (C5) yields

$$T_0^0 = V^{-2/3} \left[\frac{1}{3} V^{-2/3} (\partial_\eta \chi)^2 + \frac{1}{6} \sum_{i=1}^{3} a_i^{-2} (\partial_i \chi)^2 - \frac{1}{3} \sum_{i=1}^{3} a_i^{-2} \chi \partial_i^2 \chi \frac{1}{2} (m^2 - V^{-2/3} Q) \chi^2 \right] \tag{C6a}$$

and

$$-T_i^i = V^{-2/3} \left[\frac{1}{6} V^{-2/3} (\partial_\eta \chi)^2 + \frac{1}{3} V^{-2/3} \left(\frac{a_i'}{a_i} - \frac{1}{3} \frac{V'}{V} \right) \chi \partial_\eta \chi + \frac{2}{3} a_i^{-2} (\partial_i \chi)^2 - \frac{1}{6} \sum_{j=1}^{3} a_j^{-2} (\partial_j \chi)^2 - \frac{1}{3} a_i^{-2} \chi \partial_i^2 \chi \times \frac{1}{6} \left\{ -m^2 + V^{-2/3} \partial_\eta \left(\frac{a_i'}{a_i} - \frac{1}{3} \frac{V'}{V} \right) - V^{-2/3} Q \right\} \chi^2 \right], \tag{C6b}$$

where $\chi = V^{1/3} \phi$, $\eta = \int^t V^{-1/3} dt'$ and prime $(')$ denotes differentiation w.r.t. η and

$$Q = \frac{1}{12} \sum_{i<j} \left(\frac{a_i'}{a_i} - \frac{a_j'}{a_j} \right)^2. \tag{C6c}$$

The field equation (C3) is written as

$$\partial_\eta^2 \chi - V^{2/3} \sum_{i=1}^{3} a_i^{-2} \chi \partial_i^2 \chi + (V^{2/3} m^2 + Q) \chi = 0. \tag{C7}$$

Equation (C7) is a linear equation, so its solution can be written as

$$\chi = (2\pi)^{-3/2} \int d^3k [A_{\mathbf{k}} \chi_{\mathbf{k}}(\eta) e^{i\mathbf{k}\cdot\mathbf{x}} + A_{\mathbf{k}}^\dagger \chi_{\mathbf{k}}^*(\eta) e^{-i\mathbf{k}\cdot\mathbf{x}}], \tag{C8}$$

where $\chi_{\mathbf{k}}(\eta)$ satisfies the equation

$$\chi_{\mathbf{k}}'' + (\Omega_{\mathbf{k}}^2 + Q) \chi_{\mathbf{k}} = 0 \tag{C9a}$$

with

$$\Omega_{\mathbf{k}}^2 = V^{2/3}\omega_{\mathbf{k}}^2 = V^{2/3}\left(\sum \frac{k_i^2}{a_i^2} + m^2\right). \tag{C.9b}$$

If

$$\chi_{\mathbf{k}}^{*\prime}\chi_{\mathbf{k}} - \chi_{\mathbf{k}}^{*}\chi_{\mathbf{k}}^{\prime} = i, \tag{C10}$$

using (C4)

$$[A_{\mathbf{k}}, A_{\mathbf{k}'}] = [A_{\mathbf{k}}^{\dagger}, A_{\mathbf{k}'}^{\dagger}] = 0, \tag{C.11a}$$

$$[A_{\mathbf{k}}, A_{\mathbf{k}'}^{\dagger}] = \delta(\mathbf{k} - \mathbf{k}'). \tag{C.11b}$$

The vacuum state $|0_A\rangle$ is defined as

$$A_{\mathbf{k}}|0_A\rangle = 0. \tag{C12}$$

Connecting (C6), (C8) and (C9), one obtains

$$\rho_0 \equiv \langle 0_A|T_0^0|0_A\rangle = (16\pi^2)^{-1}V^{-4/3}\int \mathrm{d}^3k[|\chi_{\mathbf{k}}'|^2 + (\Omega_{\mathbf{k}}^2 + Q)|\chi_{\mathbf{k}}|^2] \tag{C13a}$$

and

$$\begin{aligned}(P_i)_0 &\equiv -\langle 0_A|T_i^i|0_A\rangle \\ &= \frac{1}{3}(16\pi^2)^{-1}V^{-4/3}\int \mathrm{d}^3k\left[|\chi_{\mathbf{k}}'|^2 + \left(\frac{a_i'}{a_i} - \frac{1}{3}\frac{V'}{V}\right)|\chi_{\mathbf{k}}|^2\right. \\ &\quad \left. + \left[6V^{2/3}a_i^2k_i^2 + \partial_\eta\left(\frac{a_i'}{a_i} - \frac{1}{3}\frac{V'}{V}\right) - (\Omega_{\mathbf{k}}^2 + Q)\right]|\chi_{\mathbf{k}}|^2\right].\end{aligned} \tag{C13b}$$

One can easily find that ρ_0 and $(P_i)_0$ are divergent. There exist two methods to obtain suitable finite observables: (1) regularization and (2) renormalization. Regularization means replacement of divergent quantities by well-defined expressions in a manner consistent with the physical basis of the theory. In renormalization, infinities are either absorbed in physical constants such as charge and mass or cancelled by counter-terms.

Adiabatic regularization is a subtraction scheme. The essential point in this method is the identification of contributions of the vacuum state. But it is difficult to define physical vacuum in curved space-time which really corresponds to 'no particle state'. The reason is change in gravitational field which feeds energy to perturbed mode of scalar fields. In Robertson–Walker type space-times, one has special advantage of having a priviledged class of observers, called co-moving observers, who see the universe precisely isotropic. This is why this regularization method is applicable to Robertson–Walker type space-times only as four velocity for co-moving observers are given as (1,0,0,0). Still there remains one problem which is time-dependence. To get time-independent physical vacuum state (which is a must to get 'no particle state'), the universe should expand adiabatically.

To ensure adiabaticity, Parker and Fulling [54] introduced a parameter T in the scale factors $a_i(t)(i = 1, 2, 3)$ replacing t by t/T. In the limit of large T, $a_i(t/T)$ and its

derivatives will necessarily be slowly varying functions of t, because $\partial/\partial t \rightarrow T^{-1}(\partial/\partial(t/T))$. Thus T^{-1} and its powers appear in every equation. However, original equations can be obtained by putting $T = 1$. The limit $T \rightarrow \infty$ yields the static case.

This procedure requires two important conditions.

(i) The operators A_k and $A_k^\dagger$ must be annihilation and creation operators in the adiabatic limit (arbitrarily slow time variation of components of the metric tensor). Moreover, identification of physical particles should be valid up to T^4.

(ii) Regularization of infinite quantities like ρ_0 and $(P_i)_0$ should be done using mode by mode subtraction in the integrands. Expanding the integrand in powers of T^{-1}, three leading terms are sufficient to subtract to get non-divergent quantities. Thus, adiabatic regularization is a misnomer of subtraction scheme.

To carry out the programme of adiabatic regularization, higher order WKB approximation is needed which was obtained by Parker [57] through iterative process. Later on, Chakravorty [70] obtained WKB solution of the (C9a) to all orders in an explicit form. According to Chakravorty's result positive frequency solution of (C9a) can be written as

$$\chi_k(\eta) = [2V^{1/3}(\eta)W_k(\eta)]^{-1/2} \exp\left[-i\int^\eta V^{1/3}(\eta')W_k(\eta')\mathrm{d}\eta'\right], \tag{C14a}$$

where

$$W_k(\eta) = V^{-1/3}[Y(1+\epsilon_2)(1+\epsilon_4)]^{1/2} \tag{C14b}$$

with

$$\epsilon_2 = -Y^{-3/4}\partial_\eta(Y^{-1/2}\partial_\eta Y^{1/4}) \tag{C14c}$$

and

$$\epsilon_4 = -Y^{-1/2}(1+\epsilon_2)^{-3/4}\partial_\eta\{[Y(1+\epsilon_2)]^{-1/2}\partial_\eta[(1+\epsilon_2)^{1/4}]\}, \tag{C14d}$$

Y being given as

$$Y = \Omega_k^2 + Q. \tag{C14e}$$

Using definition of Ω_k^2, given by (C9b) and Y by (C14e) in (C14b) we have

$$W_k = \omega_k[1 + \epsilon_2 + Q\Omega_k^{-2} + \epsilon_2 Q\Omega_k^{-2} + \epsilon_4 + \epsilon_4 Q\Omega_k^{-2} + \epsilon_2\epsilon_4 + Q\Omega_k^{-2} + \cdots]. \tag{C15}$$

The definition of Q, given by (C6c) yields

$$Q = O(T^{-2}). \tag{C16}$$

Using these definitions

$$\epsilon_2 = 1 + O(T^{-2}) + O(T^{-4}) + \cdots,$$

$$\epsilon_4 = O(T^{-4}) + O(T^{-6}) + \cdots.$$

So, up to order 4 of T^{-1}

$$W_{\mathbf{k}} = \omega_{\mathbf{k}}[1 + \epsilon_2 + \epsilon_4]^{1/2}, \tag{C17a}$$

where

$$\epsilon_2 = \underline{\epsilon}_2 + Q\Omega^{-2} \tag{C17b}$$

and

$$\epsilon_4 = \underline{\epsilon}_4 + \underline{\epsilon}_2 Q\Omega^{-2}. \tag{C17c}$$

Defining $\epsilon_{n(m)}$ as a term of ϵ_n which is of the order T^{-m}, one obtains up to order T^{-4},

$$|\chi|^2 = (2\Omega)^{-1}\left\{1 - \frac{1}{2}\epsilon_{2(2)} + \frac{3}{8}[\epsilon_{2(2)}]^2 - \frac{1}{2}\epsilon_{2(4)} - \frac{1}{2}\epsilon_{4(4)}\right\}, \tag{C18a}$$

$$|\chi'|^2 = (2\Omega)^{-1}\left\{\Omega^2 + \left[\frac{1}{2}\Omega^2 + \frac{1}{4}\left(\frac{\Omega'}{\Omega}\right)^2\right] + \frac{1}{2}\Omega^2\epsilon_{2(4)} + \frac{1}{2}\Omega^2\epsilon_{4(4)} - \frac{1}{8}\Omega^2[\epsilon_{2(2)}]^2 - \frac{1}{8}\left(\frac{\Omega'}{\Omega}\right)^2\epsilon_{2(2)} + \frac{1}{4}\left(\frac{\Omega'}{\Omega}\right)\epsilon'_{2(3)}\right\}, \tag{C18b}$$

$$\partial_\eta|\chi|^2 = -(2\Omega)^{-1}\left\{\left(\frac{\Omega'}{\Omega}\right) + \frac{1}{2}\left[\epsilon'_{2(3)} - \left(\frac{\Omega'}{\Omega}\right)\epsilon_{3(2)}\right]\right\}. \tag{C18c}$$

Using results, given by (C18), in (C13) divergent terms in ρ_0 and $(P_i)_0$ are obtained as

$$(\rho_0)_{\text{div.}} = (32\pi^3)^{-1}V^{-4/3}\int \mathrm{d}^3k\Omega^{-1}\left\{2\Omega^2 + \left[\frac{1}{4}\left(\frac{\Omega'}{\Omega}\right)^2 - Q\right] - \frac{1}{8}\left(\frac{\Omega'}{\Omega}\right)^2\epsilon_{2(3)} + \frac{1}{4}\left(\frac{\Omega'}{\Omega}\right)\epsilon'_{2(3)} + \frac{1}{4}\Omega^2\epsilon^2_{2(3)} + \frac{1}{2}Q\epsilon_{2(2)}\right\}, \tag{C19a}$$

$$(P_{i0})_{\text{div.}} = (96\pi^3)^{-1}V^{-4/3}\int \mathrm{d}^3k\Omega^{-1}\left\{6V^{2/3}\left(\frac{k_i}{a_i}\right)^2 + \left[\epsilon_{2(2)}\left(\Omega^2 - 3V^{2/3}\left(\frac{k_i}{a_i}\right)^2 + \left[\frac{1}{4}\left(\frac{\Omega'}{\Omega}\right) - \left(\frac{\Omega'}{\Omega}\right)\left(\frac{a'_i}{a_i} - \frac{V'}{3V}\right) + \partial_\eta\left(\frac{a'_i}{a_i} - \frac{V'}{3V}\right) - Q\right]\right.\right.$$
$$+ \Omega^2\left(\epsilon_{2(4)} + \epsilon_{4(4)} - \frac{1}{2}\epsilon^2_{2(2)}\right) - 3V^{2/3}\left(\frac{k_i}{a_i}\right)^2\left(\epsilon_{2(4)} + \epsilon_{4(4)} - \frac{3}{4}\epsilon^2_{2(2)}\right)$$
$$\left.\left. - \frac{1}{2}\epsilon_{2(2)}\left[\left(\frac{\Omega'}{2\Omega}\right)^2 - \frac{\Omega'}{\Omega}\left(\frac{a'_i}{a_i} - \frac{V'}{3V}\right) + \partial_\eta\left(\frac{a'_i}{a_i} - \frac{V'}{3V}\right) - Q\right]\right.$$
$$\left. - \frac{1}{2}\epsilon'_{2(3)}\left(\frac{a'_i}{a_i} - \frac{V'}{3V} - \frac{\Omega'}{2\Omega}\right)\right\}. \tag{C19b}$$

Regularized $(\rho_0)_{\text{reg}}$ and $(P_{i0})_{\text{reg}}$ are obtained by subtracting these divergent terms from ρ_0 and $(P_i)_0$ respectively i.e.

$$(\rho_0)_{\text{reg}} = \rho_0 - (\rho_0)_{\text{div.}}$$

and

$$(P_{i0})_{\text{reg}} = (P_i)_0 - (P_{i0})_{\text{div.}}.$$

Appendix D

Path integration

The Gaussian integral $\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2}$ can be evaluated as given below. Suppose that

$$I = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2}$$

Now

$$I^2 = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2} \int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{-y^2} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{d}y\, \mathrm{e}^{-(x^2+y^2)}.$$

Using $x = r\cos\theta$ and $y = r\sin\theta$

$$I^2 = \int_0^{\infty} \int_0^{2\pi} r\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{e}^{-r^2} = \pi,$$

which implies that

$$I = \sqrt{\pi}. \tag{D1}$$

Using this idea [57]

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-Q(x)} = \left(\frac{2\pi}{a}\right)^{1/2} \exp(b^2/2a), \tag{D2}$$

where

$$Q(x) \equiv \frac{1}{2}\, ax^2 - bx = -\frac{b^2}{2a} + \frac{1}{2}\, a\left(x - \frac{b}{a}\right)^2.$$

Taking the form of $Q(x)$ for n variables $\{x_1, x_2, \ldots, x_n\}$ as

$$\bar{Q}(x) = \frac{1}{2} \sum_{i,j=1}^{n} x^i A_{ij} x^j - \sum_{i=1}^{n} b_i x^i,$$

the generalized form of the result (D2) can be obtained. In the matrix form $\bar{Q}(x)$ can be written as

$$\bar{Q}(x) = \tfrac{1}{2} x^T A x - b^T x = \tfrac{1}{2} (x - x_0)^T A (x - x_0) - \tfrac{1}{2}\, b^T A^{-1} b, \tag{D3}$$

where A is a $n \times n$ non-singular matrix with elements A_{ij}, $x_0 = A^{-1}b, x^T$ is the transpose of x,

$$x \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and

$$b \equiv \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Now

$$\int_{-\infty}^{\infty} dx_1 dx_2 \cdots dx_n e^{-\bar{Q}(x)} = (2\pi)^{n/2} (\det A)^{-1/2} \exp(\tfrac{1}{2} b^T A^{-1} b), \tag{D4}$$

where $\det A = a_1 a_2 \cdots a_n, a_i (i = 1, 2, \ldots, n)$ being eigenvalues of the matrix A.

The effective action Γ is calculated as

$$\exp(i\Gamma/\hbar) = N \int (D\phi) \exp(i/\hbar) \left[S_0[\phi] - i \int d^4 x_E \phi(x_E) J(x_E) \right], \tag{D5}$$

where N is the normalization constant, $D\phi = (2\pi)^{-n/2} d\phi_1(x_E) \times d\phi_2(x_E) \cdots d\phi_n(x_E)$ and

$$\begin{aligned} S_0[\phi] &= \frac{1}{2} \int d^4 x \sqrt{|g|} \, [g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2] \\ &= -\frac{1}{2} \int d^4 x \sqrt{|g|} \, \phi(x) [\Box^2 - m^2] \phi(x) \\ &= \frac{i}{2} \int d^4 x_E \sqrt{|g|} \, \phi(x_E) [-\Box^2 + m^2] \phi(x_E) \end{aligned}$$

(suffix E shows Euclideanized form of the variable after Wick's rotation) and

$$\phi(x) \equiv \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix}.$$

To employ the above method of integration, Euclideanization is needed. The operator $(-\Box^2 + m^2)$ on $\phi(x_E)$ yields eigenvalues corresponding to every element of the column matrix ϕ forming a $n \times n$ diagonal matrix. So, from (D5)

$$\exp(i\Gamma/\hbar) = N (\det A)^{-1/2} \exp \frac{1}{2\hbar} [J^T (-\Box^2 + m^2) J],$$

which implies that

$$\begin{aligned} \frac{i\Gamma}{\hbar} &= \frac{1}{2\hbar} [J^T (-\Box^2 + m^2) J] + \ln [N (\det A)^{-1/2}] \\ &= \frac{1}{2\hbar} [J^T (-\Box_E^2 + m^2) J] - \frac{N}{2} \operatorname{tr} \ln A. \end{aligned} \tag{D6}$$

Appendix E

Cohomology, Betti numbers and Hodge's decomposition theorem

It is possible to distinguish between 2-dimensional non-homeomorphic spaces (two spaces are called homeomorphic if both are deformable into each other without cutting. Technically speaking, two spaces S_1 and S_2 are homeomorphic if there exists a map $f : S_1 \to S_2$ such that (1) f is one-one and onto (2) f and f^{-1} are continuous) in an intuitive manner. But this method is not effective to distinguish between higher-dimensional non-homeomorphic spaces. Homology and cohomology groups serve this purpose [71].

p-dimensional subspaces of an n-dimensional space X forming abelian groups C_p with an infinite sequence

$$\cdots \xrightarrow{\partial_{p+2}} C_{p+1} \xrightarrow{\partial_{p+1}} C_p \xrightarrow{\partial_p} C_{p-1} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_0} 0$$

are called to constitute a chain complex denoted by C^*. Here ∂_p are differentials or boundary operators such that $\partial_{p-1}\partial_p = 0$ for $p \geq 1$. Moreover ∂_p is a map such that

$$\partial_p(C_p^1 + C_p^2) = (\partial_p C_p^1) + (\partial_p C_p^2).$$

The set $\operatorname{Ker}\partial_p \equiv \{c \in C_p : \partial_p c = 0\}$ forms a subgroup in C_p called a group of p-dimensional cycles. The set $\operatorname{Im}\partial_{p+1} \equiv \{c \in C_p : c = \partial_{p+1}u\}$ forming subgroup in C_p is called a group of p-dimensional boundaries. The factor group $\operatorname{Ker}\partial_p/\operatorname{Im}\partial_p$ is called p-homology group of C^*. Thus p-homology group contains p-cycles which are not boundaries. Dual space of p-homology group is called p-cohomology group. It is also called pth de Rahm cohomology denoted as $H^p(X)$.

A general space X is not necessarily a Hausdorff space (A space is Hausdorff if open sets around any two distinct points are disjoint). If X is not Hausdorff, it may contain elements with torsion as well as torsion-free elements. pth Betti number is the dimension of pth cohomology group of torsion-free elements, because only torsion-free elements form a linear space (vector space) as linearity is lost in presence of elements with torsion.

Manifold is a space which is locally Euclidean and Hausdorff. So, all elements of a manifold are torsion-free. If manifold is real and differentiable (analytical), p-forms can be defined on it. Using differentiable p-forms, pth de Rahm cohomology group $H^p(M)$ of a real manifold M^n can be defined as the quotient space of the real vector space of closed p-forms which are not exact.

A p-form ω is closed if $\mathrm{d}\omega = 0$ and it is called exact if $\omega = \mathrm{d}\alpha$, where α is $(p-1)$-form. Similarly, a p-form is called co-closed if $\delta\omega = 0$ and co-exact, if $\omega = \delta\beta$, where β is $(p+1)$-form.

The operator d is defined as

$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i,$$

where f is 0-form (scalar) and $\mathrm{d}f$ is 1-form. Thus operator d raises the form by 1 i.e. $d : F^p(M) \to F^{(p+1)}(M)$, where $F^p(M)$ is the space of p-forms. The Hodge operator is

given as $* : F^p(M) \to F^{(n-p)}(M), 0 \leq p \leq n$. If α is a p-form given as

$$\alpha = \alpha_{i_1 \cdots i_p} \mathrm{d}x^{i_1} \wedge \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_p},$$
$$*\alpha = \alpha_{i_1 \cdots i_p} \mathrm{d}x^{i_{p+1}} \wedge \mathrm{d}x^{i_{p+2}} \wedge \cdots \wedge \mathrm{d}x^{i_n}.$$

The operator δ is given by the map $\delta : F^p(M) \to F^{(p-1)}(M)$. Thus, $\delta = - * d *$. If f is a 0-form on M^3,

$$\mathrm{d}f = \frac{\partial f}{\partial x^1}\mathrm{d}x^1 + \frac{\partial f}{\partial x^2}\mathrm{d}x^2 + \frac{\partial f}{\partial x^3}\mathrm{d}x^3$$
$$\mathrm{d}f = \frac{\partial f}{\partial x^1}\mathrm{d}x^2 \wedge \mathrm{d}x^3 + \frac{\partial f}{\partial x^2}\mathrm{d}x^3 \wedge \mathrm{d}x^1 + \frac{\partial f}{\partial x^3}\mathrm{d}x^1 \wedge \mathrm{d}x^2$$
$$\mathrm{d} * \mathrm{d}f = \frac{\partial(*\,\mathrm{d}f)}{\partial x^1}\mathrm{d}x^1 + \frac{\partial(*\,\mathrm{d}f)}{\partial x^2}\mathrm{d}x^2 + \frac{\partial(*\,\mathrm{d}f)}{\partial x^3}\mathrm{d}x^3$$
$$= \left[\frac{\partial^2 f}{\partial (x^1)^2} + \frac{\partial^2 f}{\partial (x^2)^2} + \frac{\partial^2 f}{\partial (x^3)^2}\right] \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3,$$

as $\mathrm{d}x^i \wedge \mathrm{d}x^j = -\mathrm{d}x^j \wedge \mathrm{d}x^i$. Now,

$$-\delta \mathrm{d}f = * \mathrm{d} * \mathrm{d}f = \left[\frac{\partial^2 f}{\partial (x^1)^2} + \frac{\partial^2 f}{\partial (x^2)^2} + \frac{\partial^2 f}{\partial (x^3)^2}\right],$$

which is a 0-form.

A p-form is an anti-symmetric covariant tensor. A p-form is called harmonic if $\Box\omega = 0$ (where $\Box = \mathrm{d}\delta + \delta\mathrm{d}$) which implies that $\mathrm{d}\omega = 0$ and $\delta\omega = 0$. Moreover, $\mathrm{d}\omega = 0$ and $\delta\omega = 0$ imply $\Box\omega = 0$. Each cohomology class on a compact Riemannian manifold $H^p(M)$ contains a unique harmonic representative i.e.

$$H^p(M) \equiv \{\omega \in F^p(M) : \Box\omega = 0\}.$$

To be more specific, $H^p(M)$ does not contain exact and co-exact forms.

Using these results, Hodge proposed a theorem of p-forms which states that on a compact orientable Riemannian manifold M^n, $F^p(M)$ of differentiable p-forms can be decomposed into harmonic and non-harmonic forms as [72]

$$F^p(M) = \Box F^p(M) \oplus H^p(M)$$
$$= (\mathrm{d}\delta + \delta\mathrm{d})F^p(M) \oplus H^p(M)$$
$$= \mathrm{d}F^{p-1}(M) \oplus \delta F^{p+1}(M) \oplus H^p(M).$$

Here $0 \leq p \leq n$ (p is an integer) and $H^p(M)$ is finite dimensional. In other words, if ω is a p-form on a compact orientable Riemannian manifold, it has a unique representation

$$\omega = \mathrm{d}\alpha + \delta\beta + \gamma,$$

where α is a $(p-1)$-form, β is a $(p+1)$-form and γ is a harmonic p-form. Here, it is clear that $\mathrm{d}\alpha$ is an exact form, $\delta\beta$ is a co-exact form.

As discussed above, pth Betti number b_p is the dimension of finite dimensional vector space $H^p(M)$. So, if b_p is zero on M^n, p-forms can be the sum of exact and co-exact forms only.

In the language of a physicist, exact 1-form is longitudinal component and co-exact 1-form is transverse component of a vector. In tensorial notation, a co-exact 1-form component A_μ can be recognized if it satisfies $A^{;\mu}_{\mu} = 0$ and exact 1-form component as $A_\mu = \partial_\mu f(f$ is a scalar).

Appendix F

Riemann zeta function

Using binomial expansion, one can write

$$(1-x)^{-1} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n$$

for $|x| < 1$. This kind of expansion is not possible, when $|x| \geq 1$, because in this case $\sum_{n=0}^{\infty} x^n$ is divergent. The above expansion yields

$$\left(1 - \frac{1}{p^s}\right)^{-1} = \sum_{n=0}^{\infty} \left(\frac{1}{p^s}\right)^n, \tag{F1}$$

where $s > 1$ and p is a prime number. It should be noted that 1 is not a prime number. Multiplying (F1) for all prime numbers, one obtains

$$\begin{aligned}\prod_p \left(1 - \frac{1}{p^s}\right)^{-1} &= \left(1 + \frac{1}{2^s} + \frac{1}{2^{2s}} + \cdots\right)\left(1 + \frac{1}{3^s} + \frac{1}{3^{2s}} + \cdots\right)\cdots \\ &= 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots \\ &= \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad s > 1. \end{aligned} \tag{F2}$$

This identity was found by Euler who had taken s as a real number, which sets a link between the series $\sum_{n=1}^{\infty} (1/n^s)$ and prime numbers. Riemann discussed that deepest features of prime numbers can be realized taking s as a complex number with $\mathrm{Re}\, s > 1$. So, when s is complex, $\sum_{n=1}^{\infty} (1/n^s)$ is called Riemann Zeta function written as

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathrm{Re}\, s > 1. \tag{F3}$$

The series is convergent for $\mathrm{Re}\, s > 1$ only. Integral representation of this function is given as

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} \mathrm{d}t \frac{t^{s-1}}{e^t - 1}, \tag{F4}$$

showing that $\zeta(s)$ can be analytically continued and it is single valued for all s except $s = 1$. A consequence of (D4) is given as

$$\zeta(s) = 2^s \pi^{s-1} \sin\frac{\pi s}{2}\,\Gamma(1-s)\zeta(1-s),$$

$$\zeta(s-1) = 2^{1-s}\pi^{-s}\cos(1/2\pi s)\Gamma(s)\zeta(s). \tag{F5}$$

Another important result of analytical continuation of this function is given as

$$\lim_{\epsilon\to 0}\zeta(-m+\epsilon) = (-1)^m\,\frac{B_{m+1}}{m+1}, \tag{F6}$$

where m is a natural number, B_n's are Bernoulli numbers, $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $B_5 = 0\cdots$. Thus,

$$\zeta(0) = -\frac{1}{2}, \quad \zeta(-2n) = 0, \quad \zeta(1-2n) = -\frac{B_{2n}}{2n}$$

$$\zeta(2n) = -\frac{(2\pi^{2n})}{2(2n)!}\,B_{2n} \quad (n = 1,2,3,\ldots). \tag{F7}$$

References

[1] S L Glashow, *Nucl. Phys.* **22**, 579 (1961)
[2] P D B Colins, A D Martin and E J Squires, *Particle physics and cosmology* (John Wiley and Sons, 1989)
[3] H Georgi and S L Glashow, *Phys. Rev. Lett.* **32**, 438 (1974)
[4] G Nordostrom, *Phys. Z.* **13**, 1126 (1912); *Ann. d. Phys.* **40**, 872 (1913)
[5] Th Kaluza, *Sitzungshr. Preuss. Akad. Wiss. Phys. Math. Klasse* **K1**, 966 (1921)
[6] O Klein, *Nature (London)* **118**, 516 (1926)
[7] M B Green, J H Swartz and E Witten, *Suerstring theory* (Cambridge University Press, 1987) vol. I and II
[8] P Van Nieuwenhuizen, *Phys. Rep.* **68**, 190 (1981)
[9] E Cremmer, B Julia and J Scherk, *Phys. Lett.* **B76**, 409 (1978)
[10] E Cremmer and J Scherk, *Nucl. Phys.* **B108**, 409 (1976)
[11] Z Horvath, L Palla, E Cremmer and J Scherk, *Nucl. Phys.* **B127**, 57 (1977)
[12] J F Luciani, *Nucl. Phys.* **B135**, 111 (1978)
[13] W S Massey, *Algebraic topology: An introduction* (Harcourt, Brace and World Inc., 1967)
[14] S R Huggins and D J Toms, *Nucl. Phys.* **B263**, 433 (1986)
[15] C W Misner, K S Thorne and J A Wheeler, *Gravitation* (W H Freeman, Sanfrancisco, 1973)
[16] B DeWitt, Dynamical theory of fields in *Relativity, groups and topology* edited by C DeWitt and B DeWitt (Gordon and Beach, NY, 1964)
[17] R Kerner, *Ann. Inst. H. Poincare* **9**, 143 (1968)
[18] Y M Cho and P G O Freund, *Phys. Rev.* **D12**, 1711 (1975)
[19] A Salam and J Strathdee, *Ann. Phys.* **141**, 316 (1982)
[20] S Weinberg, *Phys. Lett.* **B125**, 265 (1983)
[21] A Chodos and E Myers, *Ann. Phys.* **156**, 412 (1984)
[22] C Wetterich, *Nucl. Phys.* **B222**, 20 (1983)
[23] S K Srivastava, *Int. J. Theor. Phys.* **33**, 805 (1994)
[24] E Witten, *Nucl. Phys.* **B186**, 412 (1981)
[25] P G O Freund and M A Rubin, *Phys. Lett.* **B97**, 223 (1980)
[26] P Candelas and S Weiberg, *Nucl. Phys.* **B237**, 397 (1984)
[27] D N Page, *Phys. Rev* **D28**, 2976 (1983)
[28] Y Fujii, T Inami and M Kato, *Phys. Lett.* **B181**, 772 (1983)
[29] T Appelquist and A Chodos, *Phys. Rev. Lett.* **50**, 141 (1983); *Phys. Rev.* **D28**, 772 (1983)
[30] M A Rubin and B D Rota, *Phys. Lett.* **B127**, 55 (1983)

[31] K Tsokos, *Phys. Lett.* **B126**, 451 (1983)
[32] A Chodos and S Detweiler, *Phys. Rev.* **D21**, 2167 (1980)
[33] Q Shafi and C Wetterich, *Phys. Lett.* **B129**, 387 (1983)
[34] S Randjbar-Daemi, A Salam and J Strathdee, *Phys. Lett.* **B138**, 388 (1984)
[35] D Sahdev, *Phys. Lett.* **B137**, 155 (1984); *Phys. Rev.* **D30**, 2495 (1984)
[36] R Matzner and A Mezzacappa, *Found. Phys.* **16**, 277 (1986)
[37] E J Copeland and D J Toms, *Phys. Rev.* **D32**, 1921 (1985)
[38] M Gleiser and J G Taylor, *Phys. Rev.* **D31**, 1904 (1985)
[39] M Gleiser, S Rajput and J G Taylor, *Phys. Rev.* **D30**, 756 (1984)
[40] M Rosenbaum, M Ryan, L Urrutia and R Matzner, *Phys. Rev.* **D36**, 1032 (1987)
[41] S K Srivastava, *J. Math. Phys.* **33**, 3117 (1992)
[42] S K Srivastava, *Int. J. Theor. Phys.* **34**, 125 (1995)
[43] E Alvarez and M B Gavela, *Phys. Rev. Lett.* **51**, 931 (1983)
[44] S M Barr and L S Brown, *Phys. Rev.* **D29**, 2719 (1984)
[45] E W Kolb, D Lindley and D Seckel, *Phys. Rev.* **D51**, 1207 (1984)
[46] R B Abbott, S M Barr and S D Ellis, *Phys. Rev.* **D30**, 720 (1984); *Phys. Rev.* **D31**, 673 (1985)
[47] K Maeda, *Phys. Rev.* **D30**, 2482 (1984)
[48] T Koikawa and K Maeda, *Phys. Lett.* **149**, 82 (1984)
[49] G Gilbert, B McClain and M A Rubin, *Phys. Lett.* **B142**, 931 (1984)
[50] V A Rubakov and M E Shaposhnikov, *Phys. Lett.* **B125**, 139 (1983)
[51] M Yoshimura, *Phys. Rev.* **D30**, 344 (1984)
[52] Ya B Zeldovich, *Pis'ma Zh. Eksp. Theor. Fiz.* **12**, 207 (1970) [*JETP Lett.* **12**, 307 (1970)]
[53] Ya B Zeldovich and A A Starobinsky, *Zh. Eksp. Theor. Fiz.* **61**, 617 (1971) [*Sov. Phys. JETP* **34**, 1159 (1972)]
[54] B L Hu and L Parker, *Phys. Rev.* **D17**, 933 (1978)
[55] K Maeda, *Phys. Lett.* **B138**, 269 (1984)
[56] N D Birrell and P C W Davies, *Quantum fields in curved space-time* (Cambridge University Press, 1982)
[57] S A Fulling, L Parker and B L Hu, *Phys. Rev.* **D9**, 341 (1974)
[58] E M Lifshitz, *Zh. Eksp. Theor. Fiz.* **29**, 94 (1955) [*Sov. Phys. JETP* **2**, 73 (1956)]
[59] R Balian and B Duplantier, *Ann. Phys.* **104**, 300 (1977)
[60] K Huang, *Quarks, leptons and gauge fields* (World Scientific, 1982)
[61] I L Buchbinder, S D Odinstov and O A Fonarev, *Phys. Lett.* **B245**, 365 (1990)
[62] D J Toms, *Phys. Lett.* **129**, 31 (1983)
[63] C J Isham, *Proc. R. Soc. (London)* **A362**, 383 (1978)
[64] J S Dowker and R Banach, *J. Phys.* **A11**, 2255 (1978)
[65] D G C McKeon and T N Sherry, *Phys. Rev.* **D35**, 3584 (1987)
[66] R B Mann, L Tarasov, D G C McKeon and T Stelle, *Nucl. Phys.* **B311**, 630 (1988/89)
[67] S K Srivastava, *Class. Quantum Gravit.* **10**, 2307 (1993)
[68] S K Srivastava, *Int. J. Theor. Phys.* **35**, 171 (1996)
[69] G Gilbert and B McClain, *Nucl. Phys.* **B244**, 173 (1984)
[70] B Chakravorty, *J. Math. Phys.* **14**, 188 (1973)
[71] Yu Borisovich, N Bliznyokov, Yo Izroilevich and T Fomenko, *Introduction to topology* (Mir Publishers, Moscow, 1985)
[72] Von Westenholz, *Differential forms in mathematical physics* (North Holland, Amsterdam, 1978)

The algebra and geometry of $SU(3)$ matrices

K S MALLESH and N MUKUNDA*
Department of Studies in Physics, University of Mysore, Mysore 570 006, India
*Centre for Theoretical Studies and Department of Physics, Indian Institute of Science, Bangalore 560 012, India
*Also at Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064, India

MS received 13 March 1997

Abstract. We give an elementary treatment of the defining representation and Lie algebra of the three-dimensional unitary unimodular group $SU(3)$. The geometrical properties of the Lie algebra, which is an eight dimensional real linear vector space, are developed in an $SU(3)$ covariant manner. The f and d symbols of $SU(3)$ lead to two ways of 'multiplying' two vectors to produce a third, and several useful geometric and algebraic identities are derived. The axis-angle parametrization of $SU(3)$ is developed as a generalization of that for $SU(2)$, and the specifically new features are brought out. Application to the dynamics of three-level systems is outlined.

Keywords. $SU(3)$ matrices; octet algebra; octet geometry; $SU(3)$ axis-angle parameters.

PACS Nos 02.20; 03.65

1. Introduction

The unitary unimodular group $SU(2)$ in two complex dimensions is the simplest nontrivial example of a nonabelian compact Lie group. Its many uses in physics–spin of the electron, proton, neutron,..., isotopic spin of nucleons, description of two-level atoms and two-level quantum systems in general are very well known. At the same time its adjoint representation coincides with the three-dimensional real proper rotation group $SO(3)$, with its associated concepts of three dimensional vectors in $\mathcal{R}^3$ and their algebra. The Pauli matrices σ_j, $j = 1, 2, 3$ mediate in a natural way between the defining two-dimensional and the adjoint three-dimensional representations. Being the generators of the defining representation, the expression of a finite $SU(2)$ element as the exponential of a generator in closed form is also well known. Thus one has the familiar collection of results:

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\,\sigma_l,$$
$$\{\sigma_j, \sigma_k\} = 2\delta_{jk}; \tag{1.1a}$$
$$\underline{a}\cdot\underline{\sigma}\,\underline{b}\cdot\underline{\sigma} = \underline{a}\cdot\underline{b} + i\,\underline{a}\wedge\underline{b}\cdot\underline{\sigma}; \tag{1.1b}$$
$$a(\hat{\alpha}, \theta) = \exp(i\theta\hat{\alpha}\cdot\underline{\sigma}) = \cos\theta + i\hat{\alpha}\cdot\underline{\sigma}\ \sin\theta \in SU(2),$$
$$|\hat{\alpha}| = 1, \quad 0 \le \theta \le 2\pi. \tag{1.1c}$$

Here $\underline{a}, \underline{b}$ are (real) three-dimensional vectors, ϵ_{jkl} is the Levi–Civita symbol (structure constants of $SU(2)$), and $(\hat{\alpha}, \theta)$ are axis-angle coordinates for the general element $a(\hat{\alpha}, \theta) \in SU(2)$. The two-to-one homomorphism $SU(2) \rightarrow SO(3)$ determines an element $R(a) \in SO(3)$ for each $a \in SU(2)$:

$$R(a)_{jk} = \tfrac{1}{2}\mathrm{Tr}(\sigma_j a \sigma_k a^\dagger),$$

$$R(a')R(a) = R(a'a), \quad a', a \in SU(2). \tag{1.2}$$

The next group after $SU(2)$ in the classical unitary family is the eight-dimensional group $SU(3)$ of unitary unimodular matrices in three complex dimensions. Many representation theoretic complications expected for compact Lie groups in general, but not yet seen with $SU(2)$, do show up with $SU(3)$, making it quite nontrivial in comparison. Its use in elementary particle physics [1, 2] often exploits the canonical subgroup chain $SU(3) \supset U(2) \supset SU(2) \supset U(1)$, while its use in the nuclear physics context involves the chain $SU(3) \supset SO(3) \supset SO(2)$. The description of three-level systems [3–5] in general quantum mechanics (atoms, for instance) also involves $SU(3)$.

The purpose of this paper is to present a generalization of relations of the forms (1.1, 1.2) from $SU(2)$ to $SU(3)$, bringing out the algebraic and geometric features of the eight-dimensional octet or adjoint representation of $SU(3)$. In doing so we describe the minimum and unavoidable new features in both algebraic and geometric aspects that one must accept in the $SU(2) \rightarrow SU(3)$ transition. Among real eight component vectors in $\mathcal{R}^8$, apart from the Euclidean inner product, two kinds of bilinear products of vectors leading again to vectors – one antisymmetric and the other symmetric – play important roles, and are essential in developing a formula generalizing equation (1.1c). One of our results will indeed be an axis-angle description of $SU(3)$ elements, namely a closed-form expansion of the exponential of a general matrix in the Lie algebra $\underline{SU(3)}$ of $SU(3)$ yielding a general finite $SU(3)$ matrix. This will be seen to be considerably more complicated than the $SU(2)$ result (1.1c). In general our aim is to develop useful identities which help in getting closed form expressions, and to build up geometric pictures in some situations.

The contents of this paper are arranged as follows. Section 2 recalls the definition of the group $SU(3)$ and the generators – the λ-matrices – in the defining representation. From their commutation and anticommutation relations the structure constants f_{rst} and symmetric invariant tensor d_{rst} can be read off. Their independent nonzero components are listed. Section 3 discusses the eight-dimensional adjoint or octet representation of $SU(3)$. Based on the available invariant tensors f_{rst}, d_{rst}, two kinds of vector products among octet vectors – elements of $\mathcal{R}^8$ – are defined: an antisymmetric wedge product and a symmetric star product. Both are $SU(3)$ covariant. Apart from the geometric expression of the trilinear Jacobi identity using wedge products, several other identities involving these products and the Euclidean scalar product on $\mathcal{R}^8$ are developed. In § 4 we take up a detailed analysis of the algebraic properties of a single generator matrix in the defining representation of $SU(3)$. The geometric tools of §3 are used to get convenient forms for products, inverses, powers, determinants and the minimal equation for a general three dimensional generator matrix. A convenient way of characterizing the eigenvalue spectrum of a (suitably normalized) generator matrix, and the notion of its 'rest frame' or specific diagonal form, are developed. At all stages the $SU(3)$ covariance of the

procedures is kept in view. Section 5 introduces the concept of axis and angle parameters for $SU(3)$. This is a way of describing general one-parameter subgroups in the group. The essential differences compared to $SU(2)$ are emphasized. We also obtain a closed form expression for a general element of $SU(3)$ expressed as the exponential of a generator matrix; for this the method of going to the 'rest frame' is exploited. Section 6 briefly describes the features of Hamiltonian dynamics for three level quantum systems, based on a generalization of the Bloch spin equation familiar from two level systems. Section 7 contains some concluding remarks.

2. The defining representation of $SU(3)$ and the λ-matrices

The defining representation of the group $SU(3)$ is given by [6, 7]

$$SU(3) = \{A = 3 \times 3 \text{ complex matrix } | A^\dagger A = 1,\ \det A = 1\}. \tag{2.1}$$

This is an eight-parameter compact Lie group. The $U(2)$ and $SO(3)$ subgroups are identified (up to conjugation) by [8]

$$U(2) = \left\{ A(u) = \begin{pmatrix} u & 0 \\ 0 & (\det u)^{-1} \end{pmatrix} \middle| u \in U(2) \right\} \subset SU(3); \tag{2.2a}$$

$$SO(3) = \{A = 3 \times 3 \text{ real matrix } | A^T A = 1,\ \det A = 1\} \subset SU(3). \tag{2.2b}$$

The generalization of the Pauli matrices σ_j, in a form adapted to the $U(2)$ subgroup, leads to the eight hermitian traceless generators λ_r, $r = 1, 2, \ldots, 8$ defined as follows [6, 7]:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.3}$$

These are trace-orthonormal in the sense

$$\mathrm{Tr}(\lambda_r \lambda_s) = 2\delta_{rs}, r, s = 1, 2, \ldots, 8. \tag{2.4}$$

The commutators and anticommutators among the λ's, lead to the completely anti-symmetric structure constants f_{rst} of $SU(3)$ and to the completely symmetric d-symbols (for which there are no $SU(2)$ analogues):

$$\begin{aligned} [\lambda_r, \lambda_s] &= 2i f_{rst} \lambda_t, \\ \{\lambda_r, \lambda_s\} &= \tfrac{4}{3}\delta_{rs} + 2 d_{rst} \lambda_t; \end{aligned} \tag{2.5a}$$

$$f_{123} = 1; \quad f_{456} = f_{678} = \frac{\sqrt{3}}{2};$$
$$f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} = 1/2; \tag{2.5b}$$
$$d_{118} = d_{228} = d_{338} = -d_{888} = 1/\sqrt{3};$$
$$d_{146} = d_{157} = -d_{247} = d_{256} = d_{344} = d_{355} = -d_{366} = -d_{377} = 1/2;$$
$$d_{448} = d_{558} = d_{668} = d_{778} = -1/2\sqrt{3}. \tag{2.5c}$$

(Here only the independent nonvanishing components of f_{rst} and d_{rst} are given). The product of two λ's involves three kinds of terms:

$$\lambda_r \lambda_s = \tfrac{2}{3}\delta_{rs} + (d_{rst} + i f_{rst})\lambda_t. \tag{2.6}$$

The $U(2)$ subgroup generators are λ_1, λ_2, λ_3 (for $SU(2)$) and λ_8, while those of $SO(3)$ are λ_2, λ_5 and λ_7.

3. The adjoint representation of $SU(3)$ and the geometry of octet vectors

The adjoint representation of $SU(3)$[7] arises upon conjugation of the λ's by general $A \in SU(3)$ and expressing the result in terms of the λ's. It is a faithful representation, not of $SU(3)$, but of the quotient $SU(3)/\mathcal{Z}_3$, where $\mathcal{Z}_3$ is the centre of $SU(3)$:

$$\mathcal{Z}_3 = \{A = e^{i\omega} \cdot 1 | \omega = 0, 2\pi/3, 4\pi/3\} \subset SU(3). \tag{3.1}$$

Thus we have a three-to-one homomorphism $SU(3) \to SU(3)/\mathcal{Z}_3$. Each $A \in SU(3)$ is mapped onto an eight dimensional real orthogonal matrix $D(A) = (D(A)_{rs}) \in SO(8)$, whose matrix elements are easy to calculate:

$$A \in SU(3) \to A\lambda_r A^{-1} = D(A)_{sr}\lambda_s,$$
$$D(A)_{sr} = \tfrac{1}{2}\mathrm{Tr}(\lambda_s A \lambda_r A^\dagger),$$
$$D(A')D(A) = D(A'A). \tag{3.2}$$

Thus these matrices $D(A)$ form a very small part of the full twenty-eight dimensional group $SO(8)$. In comparison, the adjoint representation of $SU(2)$ is the same as $SO(3)$.

Let us denote general real eight component vectors in $\mathcal{R}^8$ – octet vectors – by $\underline{\alpha}$, $\underline{\beta}$, $\underline{\gamma}, \ldots$. Among them we have the usual Euclidean inner product

$$\underline{\alpha} \cdot \underline{\beta} = \alpha_r \beta_r. \tag{3.3}$$

We now define two 'vector products', one an antisymmetric wedge and the other a symmetric star, both of which lead to octet vectors again:

$$\underline{\alpha}, \underline{\beta} \in \mathcal{R}^8: (\underline{\alpha} \wedge \underline{\beta})_r = f_{rst}\alpha_s \beta_t,$$
$$\underline{\alpha} \wedge \underline{\beta} = -\underline{\beta} \wedge \underline{\alpha};$$
$$(\underline{\alpha} * \underline{\beta})_r = \sqrt{3}\, d_{rst}\alpha_s \beta_t,$$
$$\underline{\alpha} * \underline{\beta} = \underline{\beta} * \underline{\alpha}. \tag{3.4}$$

The basic $SU(3)$ covariant properties of these products follow from the fact that f_{rst} and d_{rst} are invariant tensors:

$$A \in SU(3): D(A)\underline{\alpha} \wedge D(A)\underline{\beta} = D(A)(\underline{\alpha} \wedge \underline{\beta}),$$
$$D(A)\underline{\alpha} * D(A)\underline{\beta} = D(A)(\underline{\alpha} * \underline{\beta}). \tag{3.5}$$

To express the components of $\underline{\alpha} \wedge \underline{\beta}$ and $\underline{\alpha} * \underline{\beta}$ in convenient forms in terms of those of $\underline{\alpha}$ and $\underline{\beta}$, it is useful to assemble the components α_4, α_5, α_6, α_7 of $\underline{\alpha}$ into a two-component complex column vector $\psi(\underline{\alpha})$:

$$\underline{\alpha} \in \mathcal{R}^8: \psi(\underline{\alpha}) = \begin{pmatrix} \alpha_4 - i\alpha_5 \\ \alpha_6 - i\alpha_7 \end{pmatrix}. \tag{3.6}$$

(This is related to the fact that the λ's are adapted to the $U(2)$ subgroup (2.29) of $SU(3)$). Then we have these expressions for the components of $\underline{\alpha} \wedge \underline{\beta}$:

$$(\underline{\alpha} \wedge \underline{\beta})_j = \epsilon_{jkl}\, \alpha_k \beta_l + \tfrac{1}{2} \mathrm{Im}\, \psi(\underline{\beta})^\dagger \sigma_j\, \psi(\underline{\alpha});$$
$$(\underline{\alpha} \wedge \underline{\beta})_8 = \frac{\sqrt{3}}{2}\, \mathrm{Im}\, \psi(\underline{\beta})^\dagger \psi(\underline{\alpha});$$
$$\psi(\underline{\alpha} \wedge \underline{\beta}) = \frac{i}{2} \begin{pmatrix} \beta_3 + \sqrt{3}\beta_8 & \beta_1 - i\beta_2 \\ \beta_1 + i\beta_2 & -\beta_3 + \sqrt{3}\beta_8 \end{pmatrix} \psi(\underline{\alpha}) - (\underline{\alpha} \leftrightarrow \underline{\beta}). \tag{3.7}$$

Similarly for the components of $\underline{\alpha} * \underline{\beta}$ we have:

$$(\underline{\alpha} * \underline{\beta})_j = \alpha_8 \beta_j + \beta_8 \alpha_j + \frac{\sqrt{3}}{2} \mathrm{Re} \psi(\underline{\beta})^\dagger \sigma_j\, \psi(\underline{\alpha});$$
$$(\underline{\alpha} * \underline{\beta})_8 = \alpha_j \beta_j - \alpha_8 \beta_8 - \frac{1}{2} \mathrm{Re}\, \psi(\underline{\beta})^\dagger \psi(\underline{\alpha});$$
$$\psi(\underline{\alpha} * \underline{\beta}) = \frac{1}{2} \begin{pmatrix} \sqrt{3}\beta_3 - \beta_8 & \sqrt{3}(\beta_1 - i\beta_2) \\ \sqrt{3}(\beta_1 + i\beta_2) & -\sqrt{3}\beta_3 - \beta_8 \end{pmatrix} \psi(\underline{\alpha}) + (\underline{\alpha} \leftrightarrow \underline{\beta}). \tag{3.8}$$

Now we consider some cubic relations, identities involving triple vector products, with wedges and stars in various combinations. The first of these is just a statement of the Jacobi identity for the structure constants f_{rst} and involves two wedge products:

$$\underline{\alpha} \wedge (\underline{\beta} \wedge \underline{\gamma}) + \underline{\beta} \wedge (\underline{\gamma} \wedge \underline{\alpha}) + \underline{\gamma} \wedge (\underline{\alpha} \wedge \underline{\beta}) = 0. \tag{3.9}$$

Other relations arise by calculating the triple product $\underline{\alpha} \cdot \underline{\lambda}\, \underline{\beta} \cdot \underline{\lambda} \underline{\gamma} \cdot \underline{\lambda}$ in two ways and comparing the results. We have the equality

$$\begin{aligned} \underline{\alpha} \cdot \underline{\lambda} \underline{\beta} \cdot \underline{\lambda} \underline{\gamma} \cdot \underline{\lambda} &= \frac{2}{3\sqrt{3}} \underline{\alpha} * \underline{\beta} \cdot \underline{\gamma} + \frac{2i}{3} \underline{\alpha} \wedge \underline{\beta} \cdot \underline{\gamma} + \left[\frac{2}{3} \underline{\alpha} \cdot \underline{\beta}\, \underline{\gamma} + \frac{1}{3} (\underline{\alpha} * \underline{\beta}) * \underline{\gamma} \right. \\ &\quad \left. - (\underline{\alpha} \wedge \underline{\beta}) \wedge \underline{\gamma} + \frac{i}{\sqrt{3}} ((\underline{\alpha} \wedge \underline{\beta}) * \underline{\gamma} + (\underline{\alpha} * \underline{\beta}) \wedge \underline{\gamma}) \right] \cdot \underline{\lambda} \\ &= \frac{2}{3\sqrt{3}} \underline{\alpha} \cdot \underline{\beta} * \underline{\gamma} + \frac{2i}{3} \underline{\alpha} \cdot \underline{\beta} \wedge \underline{\gamma} + \left[\frac{2}{3} \underline{\beta} \cdot \underline{\gamma} \underline{\alpha} + \frac{1}{3} \underline{\alpha} * (\underline{\beta} * \underline{\gamma}) \right. \\ &\quad \left. - \underline{\alpha} \wedge (\underline{\beta} \wedge \underline{\gamma}) + \frac{i}{\sqrt{3}} (\underline{\alpha} \wedge (\underline{\beta} * \underline{\gamma}) + \underline{\alpha} * (\underline{\beta} \wedge \underline{\gamma})) \right] \cdot \underline{\lambda}. \end{aligned} \tag{3.10}$$

The $\underline{\lambda}$-independent terms lead to the obvious results

$$\underline{\alpha} \cdot \underline{\beta}_{\wedge} \underline{\gamma} = \underline{\alpha}_{\wedge} \underline{\beta} \cdot \underline{\gamma}, \tag{3.11a}$$

$$\underline{\alpha} \cdot \underline{\beta} * \underline{\gamma} = \underline{\alpha} * \underline{\beta} \cdot \underline{\gamma}, \tag{3.11b}$$

similar to properties of the triple scalar product of vectors in $\mathcal{R}^3$. The $\underline{\lambda}$-dependent terms, upon separation of real and imaginary parts and some rearrangement, lead to further relations at the vector level:

$$\underline{\alpha} * (\underline{\beta} * \underline{\gamma}) - \underline{\gamma} * (\underline{\beta} * \underline{\alpha}) = 2(\underline{\alpha} \cdot \underline{\beta}\, \underline{\gamma} - \underline{\gamma} \cdot \underline{\beta}\underline{\alpha}) + 3\underline{\beta}_{\wedge}(\underline{\alpha}_{\wedge}\underline{\gamma}), \tag{3.12a}$$

$$\underline{\alpha}_{\wedge}(\underline{\beta} * \underline{\gamma}) + \underline{\gamma}_{\wedge}(\underline{\beta} * \underline{\alpha}) = \underline{\alpha} * (\underline{\gamma}_{\wedge}\underline{\beta}) + \underline{\gamma} * (\underline{\alpha}_{\wedge}\underline{\beta}). \tag{3.12b}$$

These are respectively antisymmetric and symmetric in the pair $\underline{\alpha}$, $\underline{\gamma}$.

The relations (3.9,11,12) are the basic $SU(3)$ covariant cubic vector relations involving three independent octet vectors.

4. Algebraic relations for $SU(3)$ generator matrices

A general traceless hermitian three dimensional matrix is of the form $\underline{\alpha} \cdot \underline{\lambda}$, $\underline{\alpha} \in \mathcal{R}^8$. For a pair of such matrices we have from eq. (2.6) the product (and square) rules

$$\underline{\alpha} \cdot \underline{\lambda} \underline{\beta} \cdot \underline{\lambda} = \frac{2}{3}\, \underline{\alpha} \cdot \underline{\beta} + \frac{1}{\sqrt{3}} \underline{\alpha} * \underline{\beta} \cdot \underline{\lambda} + i \underline{\alpha}_{\wedge} \underline{\beta} \cdot \underline{\lambda},$$

$$(\underline{\alpha} \cdot \underline{\lambda})^2 = \frac{2}{3}\, \underline{\alpha}^2 + \frac{1}{\sqrt{3}} \underline{\alpha} * \underline{\alpha} \cdot \underline{\lambda}, \tag{4.1}$$

which generalize eq. (1.1b).

Now we develop the properties of a single matrix $\underline{\alpha} \cdot \underline{\lambda}$ in some detail. The determinant is easily worked out in terms of the star product:

$$\det \underline{\alpha} \cdot \underline{\lambda} = \frac{2}{3\sqrt{3}}\, \underline{\alpha} \cdot \underline{\alpha} * \underline{\alpha}. \tag{4.2}$$

If α_0 is a ninth 'scalar', from the $SU(3)$ covariance property

$$A(\alpha_0 + \underline{\alpha} \cdot \underline{\lambda})A^{-1} = \alpha_0 + \underline{\alpha}' \cdot \underline{\lambda},$$

$$\underline{\alpha}' = D(A)\underline{\alpha}, \tag{4.3}$$

and invariance of the determinant we see that we must necessarily have

$$\det(\alpha_0 + \underline{\alpha} \cdot \underline{\lambda}) = \alpha_0^3 + c\alpha_0 \underline{\alpha}^2 + \frac{2}{3\sqrt{3}} \underline{\alpha} \cdot \underline{\alpha} * \underline{\alpha}, \tag{4.4}$$

with no term quadratic in α_0, and with some constant c. Let us now diagonalize $\underline{\alpha} \cdot \underline{\lambda}$ using a suitable $SU(3)$ transformation. We shall refer to this as 'putting $\underline{\alpha} \cdot \underline{\lambda}$ into its rest frame', and will refine this notion in the sequel. Then

$$\alpha_0 + \underline{\alpha} \cdot \underline{\lambda} = \mathrm{diag}(\alpha_0 + \alpha_3 + \alpha_8/\sqrt{3}, \alpha_0 - \alpha_3 + \alpha_8/\sqrt{3}, \alpha_0 - 2\alpha_8/\sqrt{3}),$$

$$\det(\alpha_0 + \underline{\alpha} \cdot \underline{\lambda}) = \alpha_0^3 - \alpha_0(\alpha_3^2 + \alpha_8^2) + \frac{2}{\sqrt{3}} \alpha_8(\alpha_3^2 - \alpha_8^2/3). \tag{4.5}$$

This fixes $c = -1$, so we have the general relation

$$\det(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda}) = \alpha_0^3 - \alpha_0\underline{\alpha}^2 + \frac{2}{3\sqrt{3}}\underline{\alpha}\cdot\underline{\alpha}*\underline{\alpha} \tag{4.6}$$

valid in any 'frame'.

Turning to matrix inverses, and assuming α_0 is not an eigenvalue of $-\underline{\alpha}\cdot\underline{\lambda}$, again $SU(3)$ covariance dictates the general structure

$$(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda})^{-1} = (c_1\alpha_0^2 + c_2\underline{\alpha}^2 + c_3\alpha_0\underline{\alpha}\cdot\underline{\lambda} + c_4\underline{\alpha}*\underline{\alpha}\cdot\underline{\lambda})/\det(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda}), \tag{4.7}$$

for some constants $c_1, \ldots, c_4$. Using eq. (4.6) and transposing terms we have

$$(c_1\alpha_0^2 + c_2\underline{\alpha}^2 + c_3\alpha_0\underline{\alpha}\cdot\underline{\lambda} + c_4\underline{\alpha}*\underline{\alpha}\cdot\underline{\lambda})(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda}) = \det(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda}). \tag{4.8}$$

Comparing powers of α_0 gives $c_1 = 1$, $c_2 = -1/3$, $c_3 = -1$, $c_4 = 1/\sqrt{3}$ and also the relation

$$\underline{\alpha}*(\underline{\alpha}*\underline{\alpha}) = \underline{\alpha}^2\underline{\alpha} - i\sqrt{3}(\underline{\alpha}*\underline{\alpha})_\wedge\underline{\alpha} \tag{4.9}$$

which we shall understand in another way in a moment. So for matrix inverses we have the $SU(3)$ covariant result

$$(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda})^{-1} = \left(\alpha_0^2 - \frac{1}{3}\underline{\alpha}^2 - \alpha_0\underline{\alpha}\cdot\underline{\lambda} + \frac{1}{\sqrt{3}}\underline{\alpha}*\underline{\alpha}\cdot\underline{\lambda}\right)\Big/\det(\alpha_0 + \underline{\alpha}\cdot\underline{\lambda}). \tag{4.10}$$

From the determinant relation (4.6) we see that the minimal (cubic) equation for $\underline{\alpha}\cdot\underline{\lambda}$ is

$$(\underline{\alpha}\cdot\underline{\lambda})^3 = \underline{\alpha}^2\underline{\alpha}\cdot\underline{\lambda} + \frac{2}{3\sqrt{3}}\underline{\alpha}\cdot\underline{\alpha}*\underline{\alpha}. \tag{4.11}$$

If we substitute (4.1) here for $(\underline{\alpha}\cdot\underline{\lambda})^2$ and compare coefficients we get the two relations

$$\underline{\alpha}_\wedge(\underline{\alpha}*\underline{\alpha}) = 0, \tag{4.12a}$$

$$\underline{\alpha}*(\underline{\alpha}*\underline{\alpha}) = \underline{\alpha}^2\underline{\alpha}, \tag{4.12b}$$

which explain the earlier result (4.9). Incidentally the first result above is obtainable from eq. (3.12b) by setting $\underline{\alpha} = \underline{\beta} = \underline{\gamma}$. We also obtain the following useful property of octet vectors, which generalizes the result in three dimensions that $\underline{a}_\wedge\underline{b}$ vanishes only if $\underline{b}$ is parallel to $\underline{a}$:

$$\underline{\alpha}, \underline{\beta} \in \mathcal{R}^8, \underline{\alpha}_\wedge\underline{\beta} = 0 \Leftrightarrow \underline{\beta} = c_1\underline{\alpha} + c_2\underline{\alpha}*\underline{\alpha}, c_{1,2} \text{ constants}. \tag{4.13}$$

That there are no more terms here follows from the relationship of octet vectors to traceless hermitian matrices in *three* dimensions.

Let us now trace the consequences of the minimal equation (4.11) in more detail. For brevity, denote $\underline{\alpha}*\underline{\alpha}$ by $\underline{\alpha}'$ for the moment. Then eqs (4.1), (4.11) read

$$(\underline{\alpha}\cdot\underline{\lambda})^2 = \frac{2}{3}\underline{\alpha}^2 + \frac{1}{\sqrt{3}}\underline{\alpha}'\cdot\underline{\lambda},$$

$$(\underline{\alpha}' \cdot \underline{\lambda})^2 = \frac{2}{3}\underline{\alpha}'^2 + \frac{1}{\sqrt{3}}\underline{\alpha}' * \underline{\alpha}' \cdot \underline{\lambda},$$
$$(\underline{\alpha} \cdot \underline{\lambda})^3 = \underline{\alpha}^2 \underline{\alpha} \cdot \underline{\lambda} + \frac{2}{3\sqrt{3}}\underline{\alpha} \cdot \underline{\alpha}'. \tag{4.14}$$

Now we substitute the first relation here into the second and keep simplifying till we have only linear terms in $\underline{\alpha} \cdot \underline{\lambda}$ and $\underline{\alpha}' \cdot \underline{\lambda}$:

$$\begin{aligned}\frac{2}{3}\underline{\alpha}'^2 + \frac{1}{\sqrt{3}}\underline{\alpha}' * \underline{\alpha}' \cdot \underline{\lambda} &= 3\left((\underline{\alpha} \cdot \underline{\lambda})^2 - \frac{2}{3}\underline{\alpha}^2\right)^2 \\ &= 3(\underline{\alpha} \cdot \underline{\lambda})^4 - 4\underline{\alpha}^2(\underline{\alpha} \cdot \underline{\lambda})^2 + \frac{4}{3}(\underline{\alpha}^2)^2 \\ &= \frac{2}{3}(\underline{\alpha}^2)^2 + \frac{2}{\sqrt{3}}\underline{\alpha} \cdot \underline{\alpha}' \underline{\alpha} \cdot \underline{\lambda} - \frac{\underline{\alpha}^2}{\sqrt{3}}\underline{\alpha}' \cdot \underline{\lambda}. \end{aligned} \tag{4.15}$$

This implies as consequences for any $\underline{\alpha}$, upon substituting $\underline{\alpha}' = \underline{\alpha} * \underline{\alpha}$:

$$(\underline{\alpha} * \underline{\alpha})^2 = (\underline{\alpha}^2)^2, \tag{4.16a}$$
$$(\underline{\alpha} * \underline{\alpha}) * (\underline{\alpha} * \underline{\alpha}) = 2\underline{\alpha} \cdot \underline{\alpha} * \underline{\alpha}\,\underline{\alpha} - \underline{\alpha}^2 \underline{\alpha} * \underline{\alpha}. \tag{4.16b}$$

This last relation is to be contrasted with the result of taking the star product of (4.12b) with $\underline{\alpha}$, which is

$$\underline{\alpha} * (\underline{\alpha} * (\underline{\alpha} * \underline{\alpha})) = \underline{\alpha}^2 \underline{\alpha} * \underline{\alpha}. \tag{4.17}$$

We are now in a position to deal in more detail with the eigenvalue spectrum of a single generator $\underline{\alpha} \cdot \underline{\lambda}$, and in the process refine the idea of the 'rest frame' form of $\underline{\alpha} \cdot \underline{\lambda}$. Let us hereafter assume $\underline{\alpha}$ is a unit vector, $\hat{\alpha}^2 = 1$. Then the eigenvalues of $\hat{\alpha} \cdot \underline{\lambda}$ are μ_1, μ_2, μ_3 obeying

$$\mu_1 + \mu_2 + \mu_3 = 0,$$
$$\mu_1^2 + \mu_2^2 + \mu_3^2 = \mathrm{Tr}(\hat{\alpha} \cdot \underline{\lambda})^2 = 2. \tag{4.18}$$

We can easily see that they can be ordered according to $\mu_1 \geq \mu_2 \geq \mu_3$ and the ranges can be fixed as follows:

$$\mu_{1,3} = -\frac{1}{2}\mu_2 \pm \sqrt{1 - \frac{3}{4}\mu_2^2},$$
$$-\frac{2}{\sqrt{3}} \leq \mu_3 \leq -\frac{1}{\sqrt{3}} \leq \mu_2 \leq \frac{1}{\sqrt{3}} \leq \mu_1 \leq \frac{2}{\sqrt{3}}. \tag{4.19}$$

A convenient parametrization of all three eigenvalues is by an angle φ in the range $[\pi/6, \pi/2]$ as follows:

$$\mu_1 = \frac{2}{\sqrt{3}}\sin\varphi, \quad \mu_2 = \frac{2}{\sqrt{3}}\sin(\varphi + 2\pi/3), \quad \mu_3 = \frac{2}{\sqrt{3}}\sin(\varphi + 4\pi/3). \tag{4.20}$$

The three angles occurring here are in strictly non-overlapping regions. We now define the 'rest frame' of $\hat{\alpha} \cdot \underline{\lambda}$ to be that unique diagonal form in which we have

$$\hat{\alpha}\cdot\underline{\lambda} = \operatorname{diag}\frac{2}{\sqrt{3}}(\sin\varphi, \sin(\varphi+2\pi/3), \sin(\varphi+4\pi/3)),$$
$$\hat{\alpha}_3 = \cos(\varphi+4\pi/3), \hat{\alpha}_8 = -\sin(\varphi+4\pi/3),$$
$$\hat{\alpha}_1 = \hat{\alpha}_2 = \hat{\alpha}_4 = \hat{\alpha}_5 = \hat{\alpha}_6 = \hat{\alpha}_7 = 0, \quad \pi/6 \le \varphi \le \pi/2. \tag{4.21}$$

In this 'frame' for $\hat{\alpha}$, we find that $\hat{\alpha} * \hat{\alpha}$ has similar nonzero components and is by eq. (4.16a) also a unit vector, but is not in its own 'rest frame':

$$\hat{\alpha} * \hat{\alpha}\cdot\underline{\lambda} = \operatorname{diag}\frac{2}{\sqrt{3}}(\sin(2\varphi-\pi/2), \sin(2\varphi+5\pi/6), \sin(2\varphi+\pi/6)),$$
$$(\hat{\alpha} * \hat{\alpha})_3 = 2\hat{\alpha}_3\hat{\alpha}_8 = -\sin(2\varphi+2\pi/3),$$
$$(\hat{\alpha} * \hat{\alpha})_8 = \hat{\alpha}_3^2 - \hat{\alpha}_8^2 = \cos(2\varphi+2\pi/3),$$
$$(\hat{\alpha} * \hat{\alpha})_r = 0 \quad \text{for } r = 1,2,4,5,6,7. \tag{4.22}$$

The angle φ can be inferred from the value of the invariant $\xi = \hat{\alpha}\cdot\hat{\alpha} * \hat{\alpha} = -\sin 3\varphi$: since $-1 \le \xi \le 1$ and $\pi/2 \le 3\varphi \le 3\pi/2$, ξ fixes the value of φ uniquely.

5. Axis-angle parameters and finite elements of *SU*(3)

In this section we derive the $SU(3)$ analogue of eq. (1.1c) for $SU(2)$. It is easily seen, for instance by going to the diagonal form, that every $A \in SU(3)$ can be obtained by exponentiating a suitable traceless antihermitian matrix. We now compute in closed form the matrix

$$A(\hat{\alpha},\theta) = \exp(i\theta\hat{\alpha}\cdot\underline{\lambda}) \tag{5.1}$$

set up in analogy to eq. (1.1c) for $SU(2)$. The unit octet vector $\hat{\alpha}$ is the axis, and θ is the angle, for the element $A(\hat{\alpha},\theta)$. The range for θ is discussed below. The sole $SU(3)$ scalar we can form from $\hat{\alpha}$ is the angle φ given by

$$\xi = \hat{\alpha}\cdot\hat{\alpha} * \hat{\alpha} = -\sin 3\varphi,$$
$$\pi/6 \le \varphi \le \pi/2. \tag{5.2}$$

Upon expanding the exponential in eq. (5.1) and using (4.1, 4.12b), we see that the only terms that arise are multiples of the unit matrix, of $\hat{\alpha}\cdot\underline{\lambda}$ and $\hat{\alpha} * \hat{\alpha}\cdot\underline{\lambda}$. We therefore write

$$A(\hat{\alpha},\theta) = \frac{2}{\sqrt{3}}c(\theta,\varphi) + a(\theta,\varphi)\hat{\alpha}\cdot\underline{\lambda} + b(\theta,\varphi)\hat{\alpha} * \hat{\alpha}\cdot\underline{\lambda}, \tag{5.3}$$

and proceed to determine the three $SU(3)$ scalar coefficients. For this we go to the rest frame (4.21) of $\hat{\alpha}\cdot\underline{\lambda}$ – then $\hat{\alpha} * \hat{\alpha}\cdot\underline{\lambda}$ is also diagonal (cf. eq. (4.22)) and so is $A(\hat{\alpha},\theta)$. Equation (5.3) reduces to

$$\Delta(\varphi)\begin{pmatrix} a(\theta,\varphi) \\ b(\theta,\varphi) \\ c(\theta,\varphi)\end{pmatrix} = \frac{\sqrt{3}}{2}\begin{pmatrix} \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin\varphi\right) \\ \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+2\pi/3)\right) \\ \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+4\pi/3)\right)\end{pmatrix},$$

$$\Delta(\varphi) = \begin{pmatrix} \sin\varphi & \sin(2\varphi - \pi/2) & 1 \\ \sin(\varphi + 2\pi/3) & \sin(2\varphi + 5\pi/6) & 1 \\ \sin(\varphi + 4\pi/3) & \sin(2\varphi + \pi/6) & 1 \end{pmatrix}. \tag{5.4}$$

The determinant of $\Delta(\varphi)$ is

$$\det \Delta(\varphi) = \frac{3\sqrt{3}}{2}\cos 3\varphi, \tag{5.5}$$

so $\Delta(\varphi)$ is nonsingular for $\pi/6 < \varphi < \pi/2$. In this range we have

$$\Delta(\varphi)^{-1} = \frac{2}{3\cos 3\varphi} \begin{pmatrix} -\sin 2\varphi & \sin(2\varphi + \pi/3) & \sin(2\varphi - \pi/3) \\ -\cos\varphi & \sin(\varphi + \pi/6) & -\sin(\varphi - \pi/6) \\ \frac{1}{2}\cos 3\varphi & \frac{1}{2}\cos 3\varphi & \frac{1}{2}\cos 3\varphi \end{pmatrix}, \tag{5.6}$$

and the solution for the coefficients $a(\theta,\varphi)$, $b(\theta,\varphi)$, $c(\theta,\varphi)$ is

$$a(\theta,\varphi) = \left\{ -\exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin\varphi\right)\cdot\sin 2\varphi + \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+2\pi/3)\right) \times \sin(2\varphi+\pi/3) + \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+4\pi/3)\right)\sin(2\varphi-\pi/3)\right\} \Big/ \sqrt{3}\cos 3\varphi,$$

$$b(\theta,\varphi) = \left\{ -\exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin\varphi\right)\cdot\cos\varphi + \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+2\pi/3)\right) \times \sin(\varphi+\pi/6) - \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+4\pi/3)\right)\sin(\varphi-\pi/6)\right\} \Big/ \sqrt{3}\cos 3\varphi,$$

$$c(\theta,\varphi) = \left\{ \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin\varphi\right) + \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+2\pi/3)\right) + \exp\left(\frac{2\theta}{\sqrt{3}}\cdot i\sin(\varphi+4\pi/3)\right)\right\} \Big/ 2\sqrt{3}. \tag{5.7}$$

With this the computation of $A(\hat{\alpha},\theta)$ in closed form in the generic case is complete.

The two limiting cases $\varphi = \pi/6$ and $\varphi = \pi/2$ correspond respectively to $\hat{\alpha} * \hat{\alpha} = -\hat{\alpha}$ and $\hat{\alpha} * \hat{\alpha} = \hat{\alpha}$. In these cases we find after some algebra:

$$\varphi = \pi/6 : A(\hat{\alpha},\theta) = \frac{2}{\sqrt{3}}c(\theta,\pi/6) + a(\theta,\pi/6)\hat{\alpha}\cdot\underline{\lambda},$$
$$a(\theta,\pi/6) = \{e^{i\theta/\sqrt{3}} - e^{-2i\theta/\sqrt{3}}\}/\sqrt{3},$$
$$c(\theta,\pi/6) = \{2e^{i\theta/\sqrt{3}} + e^{-2i\theta/\sqrt{3}}\}/2\sqrt{3};$$
$$\varphi = \pi/2 : A(\hat{\alpha},\theta) = \frac{2}{\sqrt{3}}c(\theta,\pi/2) + a(\theta,\pi/2)\hat{\alpha}\cdot\underline{\lambda},$$

$$a(\theta, \pi/2) = -\{e^{-i\theta/\sqrt{3}} - e^{2i\theta/\sqrt{3}}\}/\sqrt{3},$$
$$c(\theta, \pi/2) = \{2e^{-i\theta/\sqrt{3}} + e^{2i\theta/\sqrt{3}}\}/2\sqrt{3}. \tag{5.8}$$

Now we consider the question of the range of the angle variable θ. Here the situation is more intricate than in the $SU(2)$ case. With Pauli matrices, $\hat{\alpha} \cdot \underline{\sigma}$ always has eigenvalues ± 1 independent of $\hat{\alpha}$, so in eq. (1.1c) we have a fixed range $0 \leq \theta \leq 2\pi$. Equivalently, all $U(1)$ subgroups of $SU(2)$ are conjugate to one another. With $SU(3)$, however, there are infinitely many inequivalent $U(1)$ subgroups available. This means that while every $A \in SU(3)$ lies on some one-parameter subgroup and can be written as $A = A(\hat{\alpha}, \theta)$ for some $\hat{\alpha}$ and θ, the range of θ depends on $\hat{\alpha}$, more specifically on the $SU(3)$ invariant angle φ associated with $\hat{\alpha}$. The determining factor is the nature of the eigenvalues μ_1, μ_2, μ_3 of $\hat{\alpha} \cdot \underline{\lambda}$. The case $\mu_2 = 0$, $\mu_1 = -\mu_3 = 1$ leads to the range $0 \leq \theta \leq 2\pi$, which occurs for $\varphi = \pi/3$. For $\mu_2 \neq 0$, we need to distinguish between the cases of rational and irrational ratios μ_1/μ_2. For rational μ_1/μ_2, when μ_3/μ_2 is also rational, the elements $A(\hat{\alpha}, \theta)$ lie on a cyclic $U(1)$ subgroup in $SU(3)$ and θ can be taken to be in some finite interval from zero to a maximum determined by μ_2. For irrational μ_1/μ_2, when μ_3/μ_2 is also irrational, the character of the one-parameter subgroup is very different – it is the real line $\mathcal{R}$, not a cyclic $U(1)$, so $\theta \in (-\infty, \infty)$. This then is an essentially new feature with axis angle parameters for $SU(3)$ as compared to $SU(2)$.

6. Hamiltonian dynamics of three-level systems

Consider a three-level quantum system whose state is represented by a density matrix ρ. The properties

$$\rho^\dagger = \rho \geq 0, \quad \mathrm{Tr}\,\rho = 1 \tag{6.1}$$

allow us to expand ρ in terms of the λ's, bringing in a scalar c and a unit octet vector $\hat{n}$:

$$\rho = \tfrac{1}{3}(1 + c\hat{n} \cdot \underline{\lambda}),$$
$$c \leq \frac{\sqrt{3}}{2}\mathrm{cosec}(\varphi + \pi/3), \tag{6.2}$$

where the angle $\varphi \in [\pi/6, \pi/2]$ is determined by eq. (5.2): $\hat{n} \cdot \hat{n} * \hat{n} = -\sin 3\varphi$. The pure state case corresponds to $\varphi = \pi/2$ and $c = \sqrt{3}$: then $\hat{n} * \hat{n} = \hat{n}$ and the eigenvalues of ρ are (1,0,0).

Let H be a general (time-independent) Hamiltonian which we express in terms of an octet vector $\underline{h}$ and a scalar h_0:

$$H = \tfrac{1}{2}(h_0 + \underline{h} \cdot \underline{\lambda}). \tag{6.3}$$

The equation of motion for ρ (with c and $\hat{n}$ regarded as functions of time),

$$i\frac{d\rho}{dt} = [H, \rho], \tag{6.4}$$

is independent of h_0 and at first leads to

$$\dot{c}\hat{n} + c\dot{\hat{n}} = c\underline{h}_\wedge \hat{n}. \tag{6.5}$$

However since $\dot{\hat{n}}$ and $\underline{h}_\wedge\hat{n}$ are both orthogonal to $\hat{n}$, we get $\dot{c} = 0$ as expected, and the equation of motion for $\hat{n}$ becomes:

$$\dot{\hat{n}} = \underline{h}_\wedge\hat{n}. \tag{6.6}$$

This is the three level version of the Bloch equation for the spin vector familiar from two-level systems. One can now use some of the identities derived in earlier sections to verify that $\hat{n}\cdot\hat{n}*\hat{n}$ is a constant of motion:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\hat{n}\cdot\hat{n}*\hat{n} &= \dot{\hat{n}}\cdot\hat{n}*\hat{n} + 2\hat{n}\cdot\hat{n}*\dot{\hat{n}}\\ &= \underline{h}_\wedge\hat{n}\cdot(\hat{n}*\hat{n}) + 2\hat{n}\cdot\hat{n}*(\underline{h}_\wedge\hat{n})\\ &= \underline{h}\cdot\hat{n}_\wedge(\hat{n}*\hat{n}) + 2\hat{n}\cdot(\hat{n}_\wedge(\hat{n}*\underline{h}) + \underline{h}_\wedge(\hat{n}*\hat{n}))\\ &= 2\underline{h}\cdot(\hat{n}*\hat{n})_\wedge\hat{n}\\ &= 0.\end{aligned} \tag{6.7}$$

Here we used eqs (3.11a, 4.12a) and eq. (3.12b) with $\underline{\alpha} = \underline{\beta} = \hat{n}, \underline{\gamma} = \underline{h}$ to simplify terms at various stages. Thus, as expected, both the scalar c in ρ and the $SU(3)$ scalar angle φ characteristic of $\hat{n}$ are constant in time. This is consistent with the fact that $\hat{n}(t)$ evolves essentially according to the octet representation matrix $D(A(t,\hat{h}))$ where in the notation of eq. (5.1) the time t is the angle and the vector $\hat{h}$ the axis of 'rotation'.

7. Concluding remarks

We have presented a set of practical tools for carrying out calculations with finite matrices of $SU(3)$ as well as with its Lie algebra, exploiting both algebraic and geometric aspects of the situation. The space of octet vectors bears the same relation to $SU(3)$ as does ordinary Euclidean three-dimensional space to $SU(2)$. The constructions we have given for working with these vectors should prove useful in dealing with three level system dynamics. The existence and interpretation of the cubic invariant $\underline{\alpha}\cdot\underline{\alpha}*\underline{\alpha}$ and the general solution (4.13) to $\underline{\alpha}_\wedge\underline{\beta} = 0$ are noteworthy. We have also brought out the fact that in contrast to $SU(2)$, there are infinitely many distinct kinds of one-parameter subgroups in $SU(3)$, and this shows up in the axis-angle description in this case.

The main new feature in the $SU(3)$ situation, absent with $SU(2)$, is the occurrence of the d-symbols. However, for all groups $SU(n), n \geq 4$, nothing new apart from such a d-symbol is expected since one cannot in any case go beyond the commutators and anticommutators of the generators in the defining representation. It therefore is to be expected that the methods of this paper can be systematically extended to these higher dimensional groups as well.

Acknowledgements

One of us (KSM) wishes to thank the Jawaharlal Nehru Centre for Advanced Scientific Research for the award of a Visiting Fellowship, and the Centre for Theoretical Studies, Indian Institute of Science for providing facilities.

References

[1] D Griffiths, *Introduction to elementary particles* (John Wiley and Sons, New York, 1983)
[2] C Quigg, *Gauge theories of the strong, weak and electromagnetic interactions* (The Benjamin–Cummings Pub. Co., London, 1983)
[3] J N Elgin, *Phys. Lett.* **A80**, 140 (1980)
[4] P K Aravind, *J. Opt. Soc. Am.* **B3**, 1025 (1986)
[5] F T Hioe, *Phys. Rev.* **A28**, 879 (1983)
[6] M Gell-Mann and Y Neeman, *The eightfold way* (W A Benjamin Inc., New York, 1964)
[7] J J de Swart, *Rev. Mod. Phys.* **35**, 916 (1963)
[8] G Khanna, S Mukhopadhyay, R Simon and N Mukunda, *Ann. Phys. (NY)* **253**, 55 (1997)

Accurate energy-levels of the sextic-double-well oscillator $V(X) = X^6 - 3X^4$ using modified Hill determinant approach

BISWANATH RATH
Physics Department, Ravenshaw College, Cuttack 753 001, India
Address for correspondence: C/o Satyapir Kar, Sriramnagar, Khurda, Bhubaneswar 751 002, India

MS received 4 November 1996; revised 5 March 1997

Abstract. Using scaled harmonic oscillator wave functions, accurate energy levels of the sextic-double-well oscillator $V(X) = X^6 - 3X^4$ are calculated through modified Hill determinant method. Present groundstate energy remains the same as reported by Tater and Turbiner, *J. Phys.* **A26**, 697 (1993).

Keywords. Sextic-double-well oscillator; scaled wave function; groundstate energy; modified Hill determinant approach.

PACS No. 03.65

The one dimensional anharmonic oscillator potential has been used extensively in nuclear physics, particle physics, solid state physics and atomic and molecular physics. This potential has the characteristics of being a rather simple model where many non-trivial features essential to understanding quite complicated system may be implemented. Their exact solutions for arbitrary couplings are hard to find. This has culminated into the development of many fascinating approximation techniques, both perturbative [1–4] and non-perturbative [5–10]. The most commonly used oscillator is

$$H = P^2 + X^2 + \lambda X^4. \tag{1}$$

Accurate energy levels of the oscillator have been calculated [1–8] using perturbation series and Hill determinant approach. However perturbation theory has not been used successfully for double well oscillator. It is worth mentioning that original Hill determinant method of Biswas *et al* [5] has undergone many modifications for sextic well oscillator. In this context it was pointed out by Tater and Turbiner [10] that Hill determinant approach fails to yield correct groundstate energy of the sextic double well oscillator characterized by the Hamiltonian [10]

$$H = P^2 + X^6 - 3X^4. \tag{2}$$

Hence the aim of this communication is to calculate groundstate as well as excited energy levels of the sextic-double-well oscillator characterized by the Hamiltonian in (2) using modified Hill determinant approach. The procedure is as follows.

The Hamiltonian in (2) is written in terms of creation and annihilation operator using the co-ordinate transformation

$$X = \frac{(a + a^+)}{\sqrt{2W}} \tag{3a}$$

and

$$P = \sqrt{\frac{W}{2}} i(a^+ - a). \tag{3b}$$

Further parameter W is obtained by making the coefficient of a^2 or $(a^+)^2$ to zero so that the matrix element $\langle 0 | H | 2 \rangle_W$ vanishes. This procedure has been applied for a long time in nuclear physics [11] to find out the most effective Hamiltonian. Recently this procedure was suggested for parameter calculation in anharmonic oscillators either using perturbative approach [12] or non-perturbative approach [13, 14]. One can see that W satisfies the following equation

$$4W^4 - 36W - 45 = 0, \tag{4}$$

which has only one positive definite root. Now we solve the Schrödinger equation

$$H\Psi = E\Psi \tag{5}$$

by assuming

$$\Psi = \sum_{m=0}^{\infty} A_m |m\rangle_W \tag{6}$$

whose $|m\rangle_W$ are the scaled harmonic oscillator function which satisfies the relation

$$(P^2 + X^2)|m\rangle_W = (2m + 1)|m\rangle_W. \tag{7}$$

Now substituting $\Psi = \Sigma_m A_m |m\rangle_W$ in (5) we find the A_m's satisfies the following seven-term recursion relation

$$P_m A_{m-6} + Q_m A_{m-4} + R_m A_{m-2} + S_m A_m + T_m A_{m+2} + U_m A_{m+4} + V_m A_{m+6} = 0, \tag{8}$$

where

$$P_m = \frac{1}{8W^3}[m(m-1)(m-1)(m-3)(m-4)(m-5)]^{1/2}, \tag{9a}$$

$$Q_m = \left[\frac{3(2m-3)}{8W^3} - \frac{3}{4W^2}\right][m(m-1)(m-2)(m-3)]^{1/2}, \tag{9b}$$

$$R_m = \left[\frac{15(m+1)}{8W^3} - \frac{3}{W^2}\right][(m-2)[m(m-1)]^{1/2}], \tag{9c}$$

$$S_m = mW - \frac{18}{4W^2} m(m+1) + \frac{1}{8W^3}(20m^3 + 30m^2 + 40m) + \frac{W}{2} - \frac{9}{4W^2} + \frac{15}{8W^3} - E, \tag{9d}$$

Table 1. Energy levels of the sextic-double-well oscillator $V(X) = X^6 - 3X^4$.

Present calculation							
Even parity $(n = 0)$		Odd parity $(n = 1)$		Exact		Previous [10] Hill approach	
−0.660	546	0.103	079	−0.660	5	−1.672	3
3.423	833	7.586	637				
12.574	054	18.409	220				
24.989	883	32.249	273				
40.138	756	48.619	117				

Table 2. First two even states of the $V(X) = X^6 + \lambda X^2$.

λ	Present			Previous [10] Exact
3	1.935	482	1	1.935 483
	11.680	970	8	
0	1.144	802	4	1.144 802
	9.073	084	5	
−3	0.000	000	0	0
	6.298	495	9	
−11	−8.000	000	0	−8
	0.000	000	0	0

$$T_m = R_{m+2}, \tag{9e}$$

$$U_m = Q_{m+4}, \tag{9f}$$

$$U_m = P_{m+6}. \tag{9g}$$

The eigenvalue condition of the modified Hill determinant in the limit of large N is

$$\mathrm{Det}\, D_N = 0, \tag{10}$$

where

$$D_N = \begin{bmatrix} S_n & T_n & U_n & V_n & 0 & \cdots \\ R_{n+2} & S_{n+2} & T_{n+2} & U_{n+2} & V_{n+2} & \cdots \\ Q_{n+4} & R_{n+4} & S_{n+4} & T_{n+4} & U_{n+4} & \cdots \\ P_{n+6} & Q_{n+6} & R_{n+6} & S_{n+6} & T_{n+6} & \cdots \end{bmatrix},$$

and where $n = 0$ for even parity and $n = 1$ for odd parity. The zeros of D_N as a function of the parameter E give the energy eigenvalues of the problem.

The first five convergent odd levels $(n = 1)$ and even level $(n = 0)$ are tabulated in table 1. It is seen that the present groundstate energy yields the correct agreement with the groundstate energy (up to four decimals) calculated earlier by Tater and Turbiner [10]. However there are no results for comparison with the present excited values. It is worth mentioning that the previous [10] convergent groundstate energy using Hill determinant approach is −1.6723.

In this approach, the method is very simple and our choice of W reduces the Hamiltonian to its effective value. The method can be extended to any parity invariant

oscillator. Lastly we give some results for specific double well in table 2 and details will be reported elsewhere.

Acknowledgement

The author is grateful to the Director, Institute of Physics, Bhubaneswar for the hospitality during the summer where this work was completed.

References

[1] I Feranchuk, L Komarov, I Nichipor and A Ulyanenkov, *Ann. Phys. (NY)* **238**, 370 (1995) and references cited therein
[2] E J Weniger, J Cizek and F Vinette, *J. Math. Phys.* **34**, 571 (1993)
[3] E J Weniger, *Ann. Phys. (NY)* **246**, 133 (1996); *Phys. Rev. Lett.* **77**, 2859 (1996)
[4] J Cizek, E J Weniger, P Bracken and V Spirko, *Phys. Rev.* **E53**, 2925 (1996)
[5] S N Biswas, K Datta, R P Saxena, P K Srivastava and V S Varma, *J. Math. Phys.* **14**, 1190 (1973)
[6] K Banerjee, *Proc. R. Soc. London* **A364**, 265 (1978)
[7] R N Chaudhuri and M Mondal, *Pramana – J. Phys.* **37**, 13 (1991)
[8] M Znojil, *Phys. Rev.* **D26**, 3750 (1982)
[9] J Killingbeck, *Phys. Lett.* **A84**, 95 (1981); *J. Phys.* **A18**, L-1025 (1985)
[10] M Tater and A V Turbiner, *J. Phys.* **A26**, 697 (1993) (refer table 3 for numerical results using Hill-determinant approach)
[11] R K Nesbet, *Proc. R. Soc. London* **A230**, 312 (1955)
[12] B Rath, *Phys. Rev.* **A42**, 2520 (1990)
[13] R Jauregui and J Recamier, *Phys. Rev.* **A46**, 2240 (1992)
[14] F M Fernandez and E A Castro, *Phys. Rev.* **A48**, 3398 (1993)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 4,
October 1997
pp. 389–398

Hopping of chaotic dynamics in a doubly resonant parametric oscillator

JOLLY JOSE*, BIMALENDU DEB† and S DUTTA GUPTA*
*School of Physics, University of Hyderabad, Hyderabad 500 046, India
† Department of Physical Chemistry, Indian Association for the Cultivation of Science, Calcutta 700 032, India
*email: sdgsp@uohyd.ernet.in

MS received 13 March 1997

Abstract. We study a doubly resonant optical parametric oscillator where the pump can feed two pairs of signal-idler modes. We assume the presence of gain at the pump frequency. We investigate the various oscillation states of interest, namely, when only the first pair oscillates with the other pair having null amplitudes and vice versa. We demonstrate the exchange of dynamics between the mode pairs when the relevant parameters of the cavity, namely, the phase mismatch factors or the decay rates switch because of fluctuations. The exchange of dynamics is shown to be independent of the nature of dynamics, i.e. independent of whether the motion is n-periodic or chaotic. We also investigate the case where both the pairs can exhibit chaotic dynamics though these states are difficult to realize because of fluctuations.

Keywords. Parametric oscillation; chaos; switching.

PACS No. 42.65

1. Introduction

The problem of mode hopping in optical parametric oscillators [1] has attracted a lot of attention in the past few decades [2, 3]. The doubly resonant oscillator (DRO) would be the ideal source of tunable radiation, if they were without any mode hopping and cluster jumping phenomenon [4], leading to frequency instability. Very recently a dynamical model which can account for the major observed effects in DROs was presented by Agarwal and Dutta Gupta [5]. It was demonstrated that when the relevant cavity parameters (in this case they are phase mismatch factors and cavity decay rates) are changed there can be an exchange of stability between the on-states of the down-converted modes. Since any realistic system is bound to have fluctuations of the cavity parameters the dynamics has to switch from one pair of frequencies (one mode pair) to another. In fact a change in the cavity length or the refractive indices can lead to different modes spacings at the signal and idler frequencies which may result in a new pair of modes having a larger gain, thus making the new pair the preferred one. On the other hand, phase fluctuations can lead to renormalized decay rates and alter the dynamics of evolution. Phase fluctuations of the pump can be accounted for in the theory by a redefinition of the decay rates [6]. From a different angle, the physics of the

down-conversion process becomes extremely rich when there is gain at the pump frequency. It was shown by Wersinger *et al* [7] that when the pump with gain feeds only one signal-idler pair, the system can exhibit a very rich dynamics which changes from regular to chaotic (via a period-doubling bifurcation) as the decay rate of the down-converted modes increases. Besides, period-three oscillations were also observed in the system for larger values of decay rates beyond the chaotic window. Keeping the aforesaid in mind in this paper we address the question: what happens when the pump has gain and the cavity is doubly resonant? We consider the situation analogous to the case of mode hopping except that now there is an inverted medium at the pump frequency such that pump grows with a constant growth rate while propagating in the intracavity medium. To simplify the model we restrict ourselves with only two competing signal-idler pairs which are fed by the pump via a nonlinear interaction due to the presence of a $\chi^{(2)}$ material in the cavity. We show that switching of the relevant parameters (i.e., phase mismatch factor and decay rates) does lead to 'chaos hopping' for appropriate choice of other parameters. In other words, one mode pair exhibiting regular or chaotic dynamics can go to the off-state while the other starts off from the off-state to show the same response. Thus there can be an exchange of dynamics irrespective of whether it is regular or chaotic.

The organization of the paper is as follows. In § 2 we present the model and recall the results pertaining to the stability of the off-states. In § 3 we present the results of numerical integration of the dynamical system and discuss the results. Finally in § 4 we conclude the paper.

2. Mathematical formulation

We consider a pump mode with amplitude d at frequency ω_3 and two pairs of signal-idler modes, namely, a_1, a_2 and b_1, b_2, respectively. The frequencies of the modes a_1 and a_2 (b_1 and b_2) are ω_{a1} and ω_{a2} (ω_{b1} and ω_{b2}) respectively. The equations of motion for the complex amplitudes are given by

$$\dot{a}_j = -\kappa_{aj} a_j + \beta f_a \exp(-i\Delta\omega_a t) d a^*_{3-j}, \tag{1}$$

$$\dot{b}_j = -\kappa_{bj} b_j + \beta f_b \exp(-i\Delta\omega_b t) d b^*_{3-j}, \tag{2}$$

$$\dot{d} = \kappa_d d - \beta f^*_a \exp(i\Delta\omega_a t) a_1 a_2 - \beta f^*_b \exp(i\Delta\omega_b t) b_1 b_2, \tag{3}$$

where $\kappa_{aj}(\kappa_{bj})$, $j = 1, 2$, are the damping constants for the modes $a_j(b_j)$, $\Delta\omega_s = \omega_3 - \omega_{s1} - \omega_{s2}$, $(s = a, b)$ is the frequency detuning and $f_s(s = a, b)$ is the phase mismatch factor for the sth mode pair. We assumed the presence of an active medium at the pump frequency and κ_d gives the growth rate of the pump mode d. f_s, κ_{sj} and β are given by the following expressions:

$$f_s = \frac{\sin(\Delta k_s L/2)}{\Delta k_s L/2} \exp(i\Delta k_s L/2), \tag{4a}$$

$$\kappa_{sj} = \frac{\pi c}{2F_s L n_s}, \quad (s = a, b;\ j = 1, 2), \tag{4b}$$

$$\beta = \frac{2\sqrt{2} d_{\text{eff}}}{\sqrt{V}} \pi^{3/2} \sqrt{\hbar \omega_{s1} \omega_{s2} \omega_d}, \tag{4c}$$

where $\Delta k_s = k_d - k_{s1} - k_{s2}, s = a, b; F_s$ and n_s are the finesse and the refractive index at the frequency $\omega_s \cdot \kappa_{sj} (s = a, b; j = 1, 2)$ are the decay rates of the down converted modes. V and d_{eff} are the mode volume and the effective nonlinear coefficient, respectively. We assume that all the relevant frequencies are far off from any resonances of the intracavity medium. Hence we have used the same nonlinearity parameter β in eqs (1)–(3). Hereafter we will refer to $a_1, a_2 (b_1, b_2)$ as the a-(b-)pair. By making use of the following transformations

$$a_j = A_j e^{-i\Omega_{aj} t}, \quad b_j = B_j e^{-i\Omega_{bj} t}, \tag{5}$$

with

$$\Omega_{a1} + \Omega_{a2} = \Delta\omega_a, \quad \Omega_{b1} + \Omega_{b2} = \Delta\omega_b, \tag{6}$$

eqs (1)–(3) can be rewritten in the following form

$$\dot{A}_j = (-\kappa_{aj} + i\Omega_{aj}) A_j + \beta f_a d A^*_{3-j}, \tag{7}$$

$$\dot{B}_j = (-\kappa_{bj} + i\Omega_{bj}) B_j + \beta f_b d B^*_{3-j}, \tag{8}$$

$$\dot{d} = \kappa_d d - \beta f^*_a A_1 A_2 - \beta f^*_b B_1 B_2. \tag{9}$$

The set of equations (7)–(9) is analogous to the set for a doubly resonant oscillator investigated by Agarwal *et al* [5] except for the fact that in [5] the pump was assumed to decay with the decay rate κ_d and there was a drive to compensate for the decay of the pump. However, some of the features remain the same even in presence of the active medium at the pump frequency. The set of equations (7)–(9) allow for different sets of solutions, namely when one pair oscillates with the other pair in the off-state or when both the pairs oscillate. We first look at the case when, say, the a-pair is oscillating with the b-pair having null amplitudes in the long time limit. Nontrivial solutions for the a-pair is possible if and only if the pump amplitude satisfies the following relation:

$$\bar{d}^2 = \frac{\kappa_{a1}\kappa_{a2}}{\beta^2 |f_a|^2} \left[1 + \frac{\Delta\omega_a^2}{\kappa_{a1}^2 + \kappa_{a2}^2} \right] \tag{10}$$

with

$$\Omega_{aj} = \frac{\Delta\omega_a \kappa_{aj}}{\kappa_{a1} + \kappa_{a2}}. \tag{11}$$

The linear stability of the null (off) states of the b-modes can be carried out in a way analogous to that of ref. [5] leading to the following result. The inequality

$$\left| \frac{f_b}{f_a} \right| > \left(\frac{\kappa_{b1}\kappa_{b2}}{\kappa_{a1}\kappa_{a2}} \right)^{1/2} \tag{12}$$

is to be satisfied for the null state of b to be unstable. Thus for equal decay rates $(\kappa_{a1} = \kappa_{a2} = \kappa_{b1} = \kappa_{b2})$, $|f_b| > |f_a|$ implies the instability of the $B_1 = B_2 = 0$. The symmetry of eq. (12) with respect to the mode indices implies the exchange of stability (exchange of dynamics) whenever $|f_b|/|f_a|$ crosses the value unity. Similarly for a system with $|f_a| = |f_b|, \kappa_{s1} = \kappa_{s2} = \kappa_s (s = a, b)$, the inequality (12) reduces to $\kappa_b < \kappa_a$. Thus, whenever $\kappa_b < \kappa_a$ the b-mode will take off from trivial values to have a finite amplitude.

In short a mode pair with larger value of the phase mismatch factor $|f|$ or higher life time (lower decay rate) will have preference and will dominate in the long run. Note that in a realistic system fluctuations are unavoidable. Fluctuations in length or the refractive indices leads to the renormalization of the phase mismatch factor (see eq. (4a)) whereas phase or frequency fluctuations can cause a rescaling of the decay rates. Thus a random change in the physical characteristics can cause a random switching of the dynamics from one pair to the other.

It may also be noted that on-states for both the pairs are possible provided

$$\bar{d}^2 = \frac{\kappa_{a1}\kappa_{a2}}{\beta^2|f_a|^2}\left[1+\frac{\Delta\omega_a^2}{\kappa_{a1}^2+\kappa_{a2}^2}\right] = \frac{\kappa_{b1}\kappa_{b2}}{\beta^2|f_b|^2}\left[1+\frac{\Delta\omega_b^2}{\kappa_{b1}^2+\kappa_{b2}^2}\right]. \tag{13}$$

However, such states are extremely unlikely and any fluctuation leading to the violation of eq. (13) destroys the nontrivial state for one of the modes and finally the system settles down on the mode pair which has a larger $|f|$ value or a lower decay rate. Keeping in view the above analysis we now go over to the calculation of trajectories in the next section.

3. Numerical results and discussions

In this section we investigate numerically the set of equations (7)–(9). This set can be reduced to a set of real equations by introducing the real amplitudes and phases as follows

$$A_j = \bar{A}_j e^{i\varphi_{aj}}, \quad B_j = \bar{B}_j e^{i\varphi_{bj}}, \tag{14}$$

$$f_s = |f_s|e^{i\varphi_{fs}}, \quad (s = a, b), \tag{15}$$

$$d = \bar{d}e^{i\varphi_d}. \tag{16}$$

Making use of the equations (14)–(16) in (7) and (8) one can obtain the following relations:

$$\frac{d}{dt}(\bar{A}_1^2 - \bar{A}_2^2) = -2(\tilde{\kappa}_{a1}\bar{A}_1^2 - \tilde{\kappa}_{a2}\bar{A}_2^2), \tag{17}$$

$$\frac{d}{dt}(\bar{B}_1^2 - \bar{B}_2^2) = -2(\tilde{\kappa}_{b1}\bar{B}_1^2 - \tilde{\kappa}_{b2}\bar{B}_2^2), \tag{18}$$

where

$$\tilde{\kappa}_{sj} = \kappa_{sj}/\kappa_d, \quad j = 1,2; s = a, b. \tag{19}$$

For example, for $\tilde{\kappa}_{a1} = \tilde{\kappa}_{a2} = \tilde{\kappa}_a$ and $\tilde{\kappa}_{b1} = \tilde{\kappa}_{b2} = \tilde{\kappa}_b$, the quantities $(\bar{A}_1^2 - \bar{A}_2^2)$ and $(\bar{B}_1^2 - \bar{B}_2^2)$ decrease exponentially in time. Thus, if $\bar{A}_1 = \bar{A}_2 (\bar{B}_1 = \bar{B}_2)$ initially, then they remain the same for all subsequent times. Let the initial conditions be such that $\bar{A}_1 = \bar{A}_2 = \bar{A}$ and $\bar{B}_1 = \bar{B}_2 = \bar{B}$. Thus the set of seven equations (7)–(9) reduces to a set of five equations as follows

$$\dot{\bar{A}} = -\tilde{\kappa}_a\bar{A} + \tilde{\beta}|f_a|\bar{d}\bar{A}\cos(\theta_a), \tag{20}$$

$$\dot{\bar{B}} = -\tilde{\kappa}_b\bar{B} + \tilde{\beta}|f_b|\bar{d}\bar{B}\cos(\theta_b), \tag{21}$$

$$\dot{\bar{d}} = \bar{d} - \tilde{\beta}|f_a|\bar{A}^2\cos(\theta_a) - \tilde{\beta}|f_b|\bar{B}^2\cos(\theta_b), \tag{22}$$

$$\dot{\theta}_a = -\Delta\tilde{\omega}_a + \tilde{\beta}|f_a|\sin(\theta_a)(\bar{A}^2 - 2\bar{d}^2)/\bar{d} + \tilde{\beta}|f_b|\bar{B}^2\sin(\theta_b)/\bar{d}, \tag{23}$$

$$\dot{\theta}_b = -\Delta\tilde{\omega}_b + \tilde{\beta}|f_b|\sin(\theta_b)(\bar{B}^2 - 2\bar{d}^2)/\bar{d} + \tilde{\beta}|f_a|\bar{A}^2\sin(\theta_a)/\bar{d}, \tag{24}$$

where

$$\theta_a = \varphi_d + \varphi_{fa} - (\varphi_{a1} + \varphi_{a2}), \tag{25}$$

$$\theta_b = \varphi_d + \varphi_{fb} - (\varphi_{b1} + \varphi_{b2}), \quad \tilde{\beta} = \beta/\kappa_d. \tag{26}$$

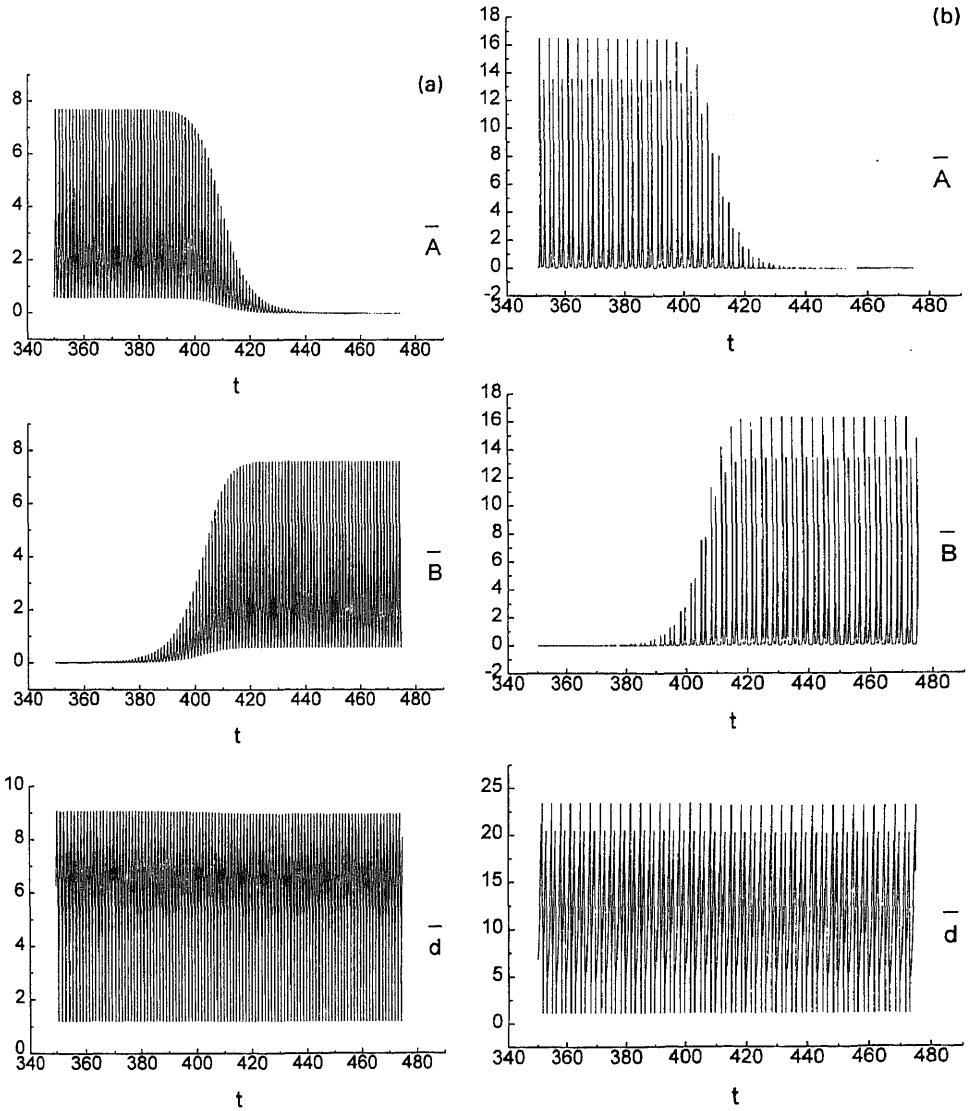


Figure 1(a) and (b).

In eqs (20)–(24) time t is scaled by κ_d. In writing eqs (20)–(24), we have used eq. (6). Note that in the absence of one of the modes a or b, the set of equations (20)–(24) reduces to eq. (1) of ref. [7], which for suitable choice of parameters exhibits regular or chaotic behavior. In the same paper a detailed analysis (including that of the Poincare section and the power spectra) for this system for $\beta = 1$ and $|f| = 1$ was carried out. It was shown that the nature of dynamics had a strong dependence on the decay rate κ of the down converted mode. For example, for $1 < \kappa \leq 3$, the motion was unbounded. In the range $3 < \kappa < 13.16$, even period bifurcations were observed. The range $13.16 \leq \kappa \leq 16.8$ was

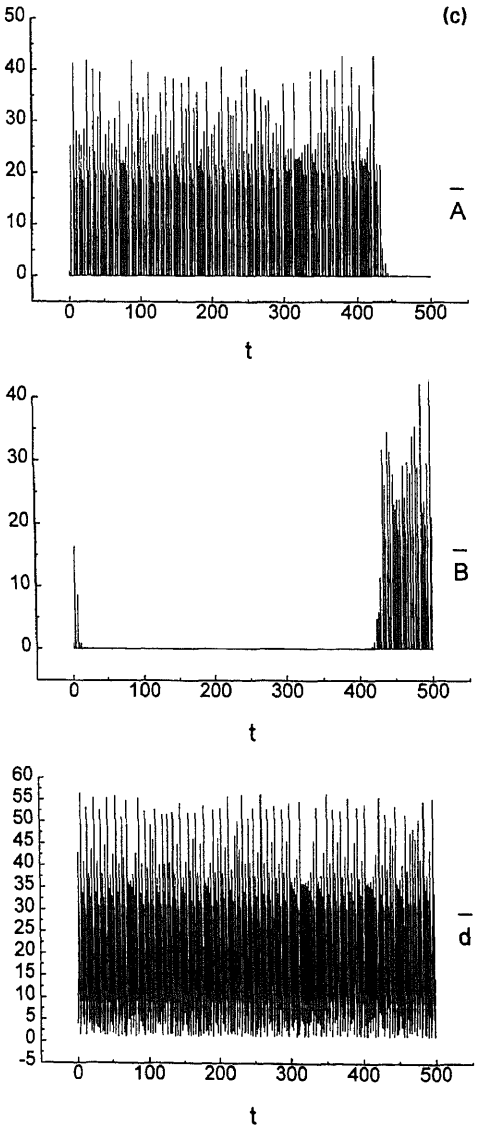


Figure 1(c).

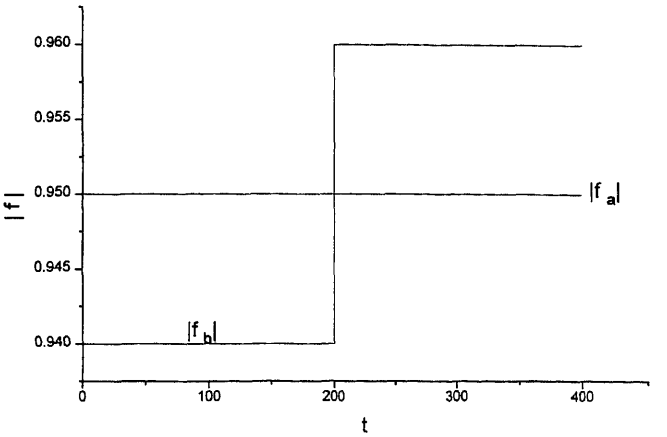


Figure 1. The temporal evolution of the mode amplitudes $\bar{A}, \bar{B}$ and $\bar{d}$ for (a) $\tilde{\kappa} = 3$, (b) $\tilde{\kappa} = 9$ and (c) $\tilde{\kappa} = 15$ when $|f_a| = 0.95$ and $|f_b|$ is switched as per eq. (27) at $t' = 200$. The other parameters are $\tilde{\beta} = 1, \delta = 0.01$ and $\Delta\tilde{\omega}_a = \Delta\tilde{\omega}_b = 2$. The inset shows $|f_a|$ and $|f_b|$ as functions of time.

characterized by lack of limit cycles and was chaotic, while larger values of κ led to period-3 bifurcation. In this paper our goal is to investigate whether the dynamics can switch from one pair to the other when the relevant parameters, namely, the phase mismatch parameter and the decay rates of the down converted modes switch at a given time, say at $t = t'$. We have carried out the numerical integration for the following two cases:

(i)
$$\tilde{\kappa}_a = \tilde{\kappa}_b = \tilde{\kappa}, |f_a| = 0.95,$$
$$|f_b| = |f_a| - \delta + \delta[1 + \text{sign}(t - t')], \tag{27}$$

(ii)
$$|f_a| = |f_b| = 1,$$
$$\tilde{\kappa}_a = \tilde{\kappa}_0\{1 + \tfrac{1}{2}[1 + \text{sign}(t - t')]\delta\}, \tag{28}$$
$$\tilde{\kappa}_b = \tilde{\kappa}_0\{1 - \delta + \tfrac{1}{2}[1 + \text{sign}(t - t')]\delta'\}, \quad \delta' = \delta^2 + 3\delta. \tag{29}$$

In eqs (28) and (29), $\tilde{\kappa}_a$ and $\tilde{\kappa}_b$ are varied such that the ratio $\tilde{\kappa}_b/\tilde{\kappa}_a$ assumes the value $1 - \delta(1 + \delta)$ before (after) switching. In all our calculations we have chosen $t' = 200$ and $\delta = 0.01, \Delta\tilde{\omega}_a = \Delta\tilde{\omega}_b = 2, \tilde{\beta} = \beta/\kappa_d = 1$. Note that such a large value of β can be achieved by a suitable choice of the pump gain, the nonlinear material or by switching to micro cavities with small quantization volume.

We now present the results for case (i). The results for the temporal evolution of the mode amplitudes $\bar{A}, \bar{B}$ and $\bar{d}$ for three values of $\tilde{\kappa}$, namely, $\tilde{\kappa} = 3, 9$ and 15 are shown in figures 1a, 1b and 1c, respectively. The inset to figure 1 shows the switching patterns for $|f_a|$ and $|f_b|$. In each of these plots we have shown the temporal evolution of the mode amplitudes. Though investigation was performed for t ranging from 0 to 600 we have shown a smaller segment in figures 1a and 1b in order to show the period one and period two oscillations. It is clear from figures 1a, 1b and 1c that as in ref. [7] the dynamics is regular for $\tilde{\kappa} = 3$ (period one), $\tilde{\kappa} = 9$ (period two) while for $\tilde{\kappa} = 15$ the motion is chaotic

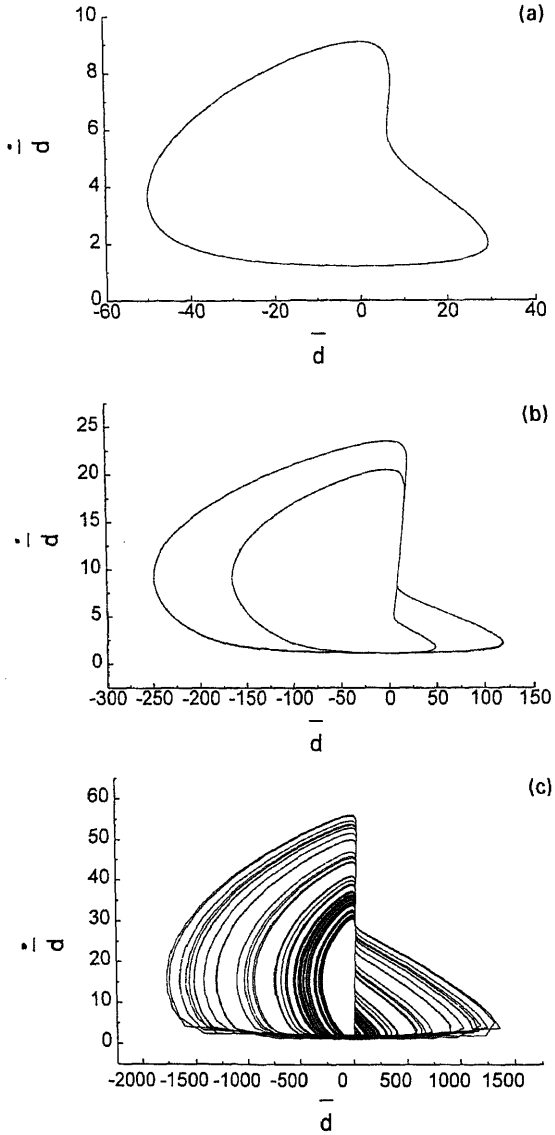


Figure 2. The phase portrait for the mode amplitudes $\bar{d}$ (i.e., $\bar{d}$ vs. $\dot{\bar{d}}$) for (**a**) $\tilde{\kappa} = 3$, (**b**) $\tilde{\kappa} = 9$ and (**c**) $\tilde{\kappa} = 15$. The other parameters are as in figure 1.

(figure 1c). This is depicted by the phase portraits of $\bar{d}$ which is shown in figures 2a, 2b and 2c. It is also clear from figure 1 that before switching, $|f_a| > |f_b|$ and the a-mode shows nonzero oscillations while the b-mode is off. The dynamics switches from a to b-mode once $|f_b|$ becomes larger than $|f_a|$. The switching of the dynamics is independent of whether the motion is with one-period, two-period or chaotic in nature. It may also be noted that the exchange of dynamics is also independent of how switching takes place.

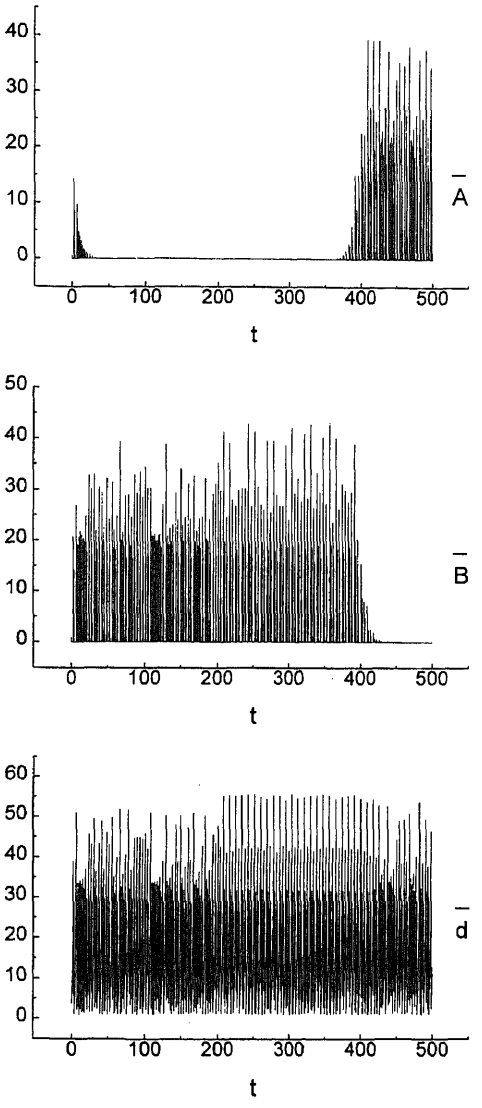


Figure 3. The temporal evolution of the mode amplitudes $\bar{A}$, $\bar{B}$ and $\bar{d}$ when $\tilde{\kappa}_a$ and $\tilde{\kappa}_b$ are switched as per (28) and (29) for $\tilde{\kappa}_0 = 15$, $|f_a| = |f_b| = 1$ and $t' = 200$. The other parameters are as in figure 1.

For example, instead of a step function (eq. (27)) one could have chosen an exponential switching.

We now look at case (ii) when $|f_a|$ and $|f_b|$ are held fixed at unity while $\tilde{\kappa}_a$ and $\tilde{\kappa}_b$ are switched at $t = t'$ such that $\tilde{\kappa}_b/\tilde{\kappa}_a$ is $1 - \delta(1 + \delta)$ before (after) switching as per eqs (28) and (29). We have performed the investigation for three values of $\tilde{\kappa}_0$, namely, $\tilde{\kappa}_0 = 3, 9$ and 15 which correspond to period one, period two and chaotic motion. We present the results only for $\tilde{\kappa}_0 = 15$ (see figure 3). It is clear from figure 3 that for the same value of

and both of them can show nontrivial oscillations (not shown). However, as mentioned earlier, on-states for both the modes are extremely unlikely. Any fluctuation leading to the violation of the second equality in eq. (13) will lead to a situation when only one mode pair survives.

4. Conclusion

We have studied a doubly resonant parametric oscillator where a pump with gain can feed two pairs of signal-idler modes. We have demonstrated both theoretically and numerically that there can be an exchange of dynamics between the down-converted mode pairs, when the relevant parameters like phase mismatch factor or cavity decay rates change due to fluctuations. The switching of dynamics from on-to-off or from off-to-on for the pairs takes place even when the dynamical system exhibits chaotic behavior. Moreover, there can be a state where both the down converted pairs exhibit chaotic motion though because of ever present fluctuations, this state is highly unlikely.

Acknowledgements

One of the authors (SDG) would like to thank G S Agarwal for initiating him to this class of problems. SDG is also thankful to the Department of Science and Technology for supporting this work.

References

[1] J A Giordmaine and R C Miller, *Phys. Rev. Lett.* **14**, 973 (1965); *Appl. Phys. Lett.* **9**, 298 (1966)

[2] For various research activities on optical parametric oscillators see, for example, the references in the feature issue (September and November) on Optical Parametric Oscillation and Amplification, *J. Opt. Soc. Am.* **B10**, 1656 (1993)

[3] For a review of the optical parametric oscillators see, for example, R L Byer in *Treatise in quantum electronics* edited by H Rabin and C L Tang (New York, Academic, 1973) p. 588–702
R J Smith, in *Advances in Lasers* edited by A K Levine and A J deMaria (New York, Dekker 1976) vol. 4
F Zernike and J E Midwinter, *Applied nonlinear optics* (New York, Wiley, 1973) p. 153–176
For recent review see, C L Tang, Optical parametric processes and inorganic nonlinear optical crystals, in *Proceedings of the International School of Physics ⟨Enrico Fermi⟩, Varena, Italy* (Amsterdam, North Holland, 1994) p. 97–129
C L Tang, *Int. J. Nonlinear Opt.* **3**, 205 (1994)

[4] For a detailed description of mode hopping see, S E Harris, *Proc. IEEE* **57**, 2096 (1969)
J Falk, *IEEE J. Quant. Electron.* **QE-7**, 230 (1971)
R J Smith, *IEEE J. Quant. Electron.* **QE-9**, 530 (1973)

[5] G S Agarwal and S Dutta Gupta, *A model for mode hopping in optical parametric oscillators J. Opt. Soc. Am.* **B14**, 2174 (1997)

[6] G S Agarwal, *Phys. Rev.* **A18**, 1490 (1978)

[7] Jean-Maire Wersinger, John M Finn and Edward Ott, *Phys. Rev. Lett.* **44**, 453 (1980)

The effect of environment induced pure decoherence on the generalized Jaynes–Cummings model

GAUTAM GANGOPADHYAY and S H LIN*
S N Bose National Centre for Basic Sciences, JD Block, Sector III, Salt Lake City, Calcutta 700 091, India
*Institute of Atomic and Molecular Sciences, Academia Sinica, P.O. Box 23-166, Taipei, Taiwan 106, R.O.C

MS received 21 September 1996; revised 5 April 1997

Abstract. We have studied the effect of environment induced pure decoherence on the generalized Jaynes–Cummings model (JCM). This generalized JCM is introduced to take into account both atom-field interaction and a class of spin-orbit interactions in the same framework. For generalized JCM with atom-field interaction, it is shown that along with the suppression of the oscillatory behaviour of the atomic and field variables, in the steady state, atomic energy is transferred to the field or vice versa through the dressed state coherence depending on the initial condition of the atom-field system and the model under consideration. It is also shown that initial Poissonian field acquires a sub-Poissonian character in the steady state and thus all the nonclassical properties are not erased by the decoherence in JCM. An interesting effect of this decoherence mechanism is that it affects the population and coherence properties of the individual subsystem in a different way. As an example of generalized JCM with spin-orbit interaction, the dynamics of spin of the hydrogen atom in a magnetic field is studied to show the effect of decoherence.

Keywords. Decoherence; Jaynes–Cummings model; generalized JCM.

PACS Nos 42.50; 03.65

1. Introduction

The Jaynes–Cummings model (JCM) is a two-state system interacting with a single mode field in the rotating wave approximation (RWA) [1]. This model was first used to examine the classical aspects of spontaneous emission but it was subsequently discovered that in the JCM atomic population histories presented direct evidence for the discreteness of photons. Also many nonclassical features of photon statistics were discovered through this model. More recently, JCM has been used to elucidate quantum correlation and the formation of macroscopic superposed quantum states. Apart from exact solutions of different variants of JCM, the remarkable advancement in cavity quantum electrodynamics experiments [2] involving single atoms within single mode cavities have made the model a very good testing ground of the quantum phenomena of the atom-field entangled system.

It is well-known that the surroundings have profound effects on the evolution of a system of interest. In particular, quantum noise due to the surroundings may result in nonunitary evolution of the system resulting in decoherence and relaxation. One can

naturally ask the question what will be the effect of decoherence on the atom and field variables which are interacting through the JCM. Here unlike the normal case of spontaneous emission we are looking at a situation where there is no energy dissipation but only a decoherence from the composite system, i.e, a case of pure decoherence. Although in many cases the effect of decoherence is studied for single systems of interest [3–5], it may be worthwhile to understand the effect of decoherence on the joint subsystems [6]. We examine the fate of the individual atom and field variables when the decoherence acts on the composite atom-field system.

In this work we have given a solution for generalized JCM undergoing environment induced pure decoherence in terms of the atom-field dressed states. By this approach one can observe that the decoherence does not affect the dressed state population but affect only the coherence of the dressed state levels. But when one looks at the individual atomic or field population it decays to a steady state and in the steady state depending on the initial condition and specific model under study, atomic energy is transferred to the field or vice-versa. This is due to the fact that the dressed state coherence has an effect on the atomic and field population and as the decoherence model is energy conserving, the total energy is distributed in this way.

We have also shown that the nonclassical properties of the field get affected by the decoherence. Although squeezing and other phase related nonclassical properties are suppressed by the decoherence, the sub-Poissonian [7] nature of the field is not completely suppressed by the decoherence. This is exemplified with some variants of atom-field interaction models.

We have pointed out that in JCM with atom-field interaction, the decoherence mechanism as proposed affects the population and coherence properties of the individual subsystem in a different way.

We have also introduced a more generalized JCM which not only accounts for the atom-field interaction but also for a class of spin-orbit interactions. As an explicit example, it is shown that a hydrogen atom in a magnetic field can also be solved. This system undergoing decoherence can also be studied using dressed states. The characteristic spin dynamics is studied for a particular value of orbital angular momentum.

The paper is organized as follows: In § 2(a) we present the master equation for pure decoherence and its effect on generalized Jaynes–Cummings (JC) model with atom-field interaction. The effect of decoherence mechanism on the transient atom-field dynamics is discussed in § 2(b). Taking a particular nonlinear JC model in § 3 we show how the nonclassical properties of the field and and atom evolves in time. In § 4 we have given a more general Hamiltonian which describes not only the atom-field interaction but also a class of spin-orbit interactions, in general. Section 5 is devoted to show how the dynamics of spin is affected by the decoherence in the case of hydrogen atom in presence of a magnetic field. In § 6 we conclude the paper.

2a. Master equation for pure decoherence and its effect on the generalized JCM with atom-field interaction

To start with, we allow an interaction between the system and its environment reservoir which gives only a decoherence in the system but no energy dissipation from the system.

$$H_R = \sum_{j=0}^{\infty} \hbar\omega_j b_j^\dagger b_j, \tag{2.2}$$

where ω_j is the frequency and $b_j^\dagger(b_j)$ is the creation (annihilation) operator for the jth reservoir mode. We consider the interaction part V as a quantum non-demolition coupling satisfying

$$[H, V] = 0. \tag{2.3}$$

Such couplings have been previously examined by others and the examples include, optical four-wave mixing [8], many particle bosonic systems [9] etc.

Equation (2.3) implies that V is a constant of motion generated by H and V can be taken as some function of H. We take for example,

$$V = \hbar H \sum_j [\kappa(\omega_j) b_j^\dagger + \kappa^*(\omega_j) b_j], \tag{2.4}$$

where $\kappa(\omega_j)$ is a c-function.

Thus one can write the Liouville equation for the system and the bath in the interaction picture (IP) as

$$\frac{\partial \eta}{\partial t} = -\frac{i}{\hbar}[V_I(t), \eta(t)], \tag{2.5}$$

where $\eta(t)$ and $V_I(t)$ are the joint system-bath density and coupling term in IP respectively and $V_I(t)$ is given by

$$V_I(t) = \exp[i(H + H_R)t/\hbar] V \exp[-i(H + H_R)t/\hbar]. \tag{2.6}$$

After integrating eq. (2.5) and iterating twice one can write

$$\eta(t) - \eta(0) = -(i/\hbar) \int_0^t [V_I(t'), \eta(0)] dt' + (-i/\hbar)^2 \int_0^t dt' \int_0^{t'} dt'' [V_I(t'), [V_I(t''), \eta(t'')]]. \tag{2.7}$$

The first term in the RHS of eq. (2.7) will not contribute due to the random initial condition of the bath which is often used. Now we make a number of approximations as follows [10].

(i) We integrate eq. (2.7) with respect to t thereby making a coarse graining approximation.

(ii) We take $\eta(t) = \eta_s(t)\eta_b(t)$ where s and b correspond to system and bath. Further assume the bath is in thermal equilibrium and thus we can write $\eta_b(t) = \eta_b$ (at equilibrium).

(iii) We use the Born–Markov approximation whereby we take $\eta_s(t') = \eta_s(t)$ i.e, truncate the series after second order and make the limits of integration up to infinity.

Now taking trace over the bath varibles one arrives at the equation of motion of the reduced system density operator $\eta_s(t)$ in IP as

$$\frac{d\eta_s}{dt} = -\left\{[HH\eta_s - 2H\eta_s H + \eta_s HH]\sum_j |\kappa(\omega_j)|^2 \int_0^\infty dt' \langle b_j b_j^\dagger\rangle \exp[i\omega_j(t-t')]\right.$$
$$\left. + [HH\eta_s - 2H\eta_s H + \eta_s HH]\sum_j |\kappa(\omega_j)|^2 \int_0^\infty dt' \langle b_j^\dagger b_j\rangle \exp[-i\omega_j(t-t')]\right\}. \tag{2.8}$$

Thus returning back to the Schrödinger picture (SP) we can write

$$\frac{d\rho}{dt} = -(i/\hbar)[H,\rho] - \left\{[HH\rho - 2H\rho H + \rho HH]\sum_j |\kappa(\omega_j)|^2\right.$$
$$\times \int_0^\infty dt' \langle b_j b_j^\dagger\rangle \exp[i\omega_j(t-t')] + [HH\rho - 2H\rho H + \rho HH]$$
$$\left.\times \sum_j |\kappa(\omega_j)|^2 \int_0^\infty dt' \langle b_j^\dagger b_j\rangle \exp[-i\omega_j(t-t')]\right\}, \tag{2.9}$$

where ρ is the reduced system density operator in SP

Now the survival of the integrals in eq. (2.9) within the Markovian approximation means that approximately zero frequency bath modes are important and very often the characteristic frequency of the system is much larger than that. Thus to attain a finite relaxation constant one needs to have the thermal average number density of bath at approximately zero frequency modes high enough which is possible only at high temperatures. A typical integral can be shown as

$$\sum_j |\kappa(\omega_j)|^2 \int_0^\infty d\tau \langle b_j b_j^\dagger\rangle \exp[i\omega_j\tau] = \sum_j \int_0^\infty d\tau \exp[i\omega_j\tau]|\kappa(\omega_j)|^2[1+\bar{n}(\omega_j)]. \tag{2.10}$$

Now converting the sum into an integral and assuming a continuous spectral density of reservoir modes, $s(\omega)$ and taking the thermal average excitation number in the bath as

$$\bar{n}(\omega_j) = [\exp(\hbar\omega_j/kT) - 1]^{-1}.$$

The RHS of eq. (2.10) can be written as

$$\text{RHS of eq.(2.10)} = (1+\bar{n}(\omega)) \lim_{\omega\to 0} \frac{2\pi|\kappa(\omega)|^2 s(\omega)}{\omega} = \gamma kT/\hbar,$$

where k is the Boltzmann constant and $\gamma = \lim_{\omega\to 0}(2\pi|\kappa(\omega)|^2 s(\omega))/\omega$ and is assumed to be finite, i.e., $\gamma > 0$. We have taken the high temperature limit i.e, $\bar{n}(\omega) \approx kT/\hbar\omega$ and thus neglecting the zero temperature term in eq. (2.9), we arrive at the desired master

equation for the reduced system as

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] - c[HH\rho + \rho HH - 2H\rho H]. \tag{2.11}$$

Here c is given by $\gamma kT/\hbar$. The frequency shift term is neglected here.

We first note that eq. (2.11) is a typical model for studying decoherence. The decoherence arising due to the assumption that there exists a classical measurement apparatus during which the measurement process destroys quantum coherences [4] has also been considered. Another approach relies on the assumption in addition to conventional quantum mechanics that in a sufficiently short time scale, the quantum system does not evolve continuously under a unitary transformation but in a stochastic sequence of identical unitary transformation [5, 11]. In that case the rate of decoherence does not depend on temperature but on the inverse of the time scale of stochastic perturbation.

In what follows we consider a two state system with levels $|b\rangle$(ground) and $|a\rangle$(excited) coupled to a single mode field. A basis for the whole Hilbert space is given as $\{|n\,b\rangle, |n\,a\rangle, n = 0, 1, 2, \ldots\}$, where $|n\,b\rangle(|n\,a\rangle)$ denotes a state with n photons with the atom in its ground (excited) state. We assume that the atomic transition is mediated by k photons; i.e, the Hamiltonian in terms of this basis is given, in the RWA, by

$$H = \sum_{n=0}^{\infty} b_n |n\,b\rangle\langle n\,b| + a_n |n\,a\rangle\langle n\,a| + \sum_{n=0}^{\infty} R_n |n\,a\rangle\langle n+k\,b| + R_n^* |n+k\,b\rangle\langle n\,a|, \tag{2.12}$$

where b_n and a_n are the energies of the state $|n\,b\rangle$ and $|n\,a\rangle$ respectively and R_n describes the coupling. This Hamiltonian is diagonalized as follows:

$$H|n\,b\rangle = b_n|n\,b\rangle, \quad n < k, \quad \text{and} \quad H|\Psi_n^{\pm}\rangle = H_n^{\pm}|\Psi_n^{\pm}\rangle, \quad n \geq k, \tag{2.13}$$

where

$$H_n^{\pm} = \frac{b_{n+k} + a_n}{2} \pm \Omega_n, \tag{2.14}$$

and

$$\Omega_n^2 = (\Delta_n/2)^2 + |R_n|^2; \tag{2.15}$$

with detuning $\Delta_n = a_n - b_{n+k}$. The eigenstates $|\Psi_n^{\pm}\rangle$ are the well-known dressed states [12]

$$\begin{aligned} |\Psi_n^+\rangle &\equiv |+n\rangle = \cos\theta_n |b\,n+k\rangle + \sin\theta_n |a\,n\rangle, \\ |\Psi_n^-\rangle &\equiv |-n\rangle = -\sin\theta_n |b\,n+k\rangle + \cos\theta_n |a\,n\rangle, \end{aligned} \tag{2.16}$$

with θ_n given as

$$\tan\theta_n = \frac{R_n}{\Delta_n/2 + \Omega_n}. \tag{2.17}$$

This structure of the Hamiltonian and its solution are valid for many variants of the atom-field interaction models. This model Hamiltonian can be further extended to describe a

hydrogen atom in an arbitrarily strong magnetic field and also for systems where the field or the atom can be replaced by a finite dimensional multilevel system which is discussed in § 4.

Here we show that the equation of motion (2.11) is exactly solvable for the Hamiltonian of generalized JC model. To solve the equation of motion we used the dressed state density matrix elements as

$$\rho_{1nm} = \langle +n|\rho|+m\rangle, \quad \rho_{4nm} = \langle -n|\rho|-m\rangle,$$
$$\rho_{2nm} = \langle +n|\rho|-m\rangle, \quad \rho_{3nm} = \langle -n|\rho|+m\rangle. \tag{2.18}$$

The population and coherence of the dressed state elements evolve as

$$\dot{\rho}_{1nn} = 0 = \dot{\rho}_{4nn}$$

and

$$\dot{\rho}_{2nm} = [-i(H_n^+ - H_m^-) - c(H_n^+ - H_m^-)^2]\rho_{2nm},$$
$$\dot{\rho}_{3nm} = [-i(H_n^- - H_m^+) - c(H_n^- - H_m^+)^2]\rho_{3nm},$$
$$\dot{\rho}_{1nm} = [-i(H_n^+ - H_m^+) - c(H_n^+ - H_m^+)^2]\rho_{1nm},$$
$$\dot{\rho}_{4nm} = [-i(H_n^- - H_m^-) - c(H_n^- - H_m^-)^2]\rho_{4nm}. \tag{2.19}$$

It is immediately apparent that the decoherence does not affect the dressed state populations but only induces a decay in the dressed state coherence. Nevertheless the decay of dressed state coherence terms depends on the free evolution of the composite system.

A variety of initial conditions of the atom-field system can be studied but here we restrict ourselves to simple initial conditions that the atom is initially in the excited state and the field is in a coherent state

$$|\alpha\rangle = \sum_{n=0}^{\infty} C_n|n\rangle, \tag{2.20}$$

where

$$C_n = \exp(-|x|^2/2)x^n/\sqrt{(n!)}, \tag{2.21}$$

with $|x|^2$ as the average number of photons. Initial conditions in terms of the dressed state elements can be written as follows:

$$\rho_{1nm}(0) = \sin\theta_n \sin\theta_m C_n C_m^*,$$
$$\rho_{4nm}(0) = \cos\theta_n \cos\theta_m C_n C_m^*,$$
$$\rho_{2nm}(0) = \sin\theta_n \cos\theta_m C_n C_m^*,$$
$$\rho_{3nm}(0) = \cos\theta_n \sin\theta_m C_n C_m^*. \tag{2.22}$$

We define the atomic density matrix as

$$\kappa = \mathrm{Tr}_f[\rho] \tag{2.23}$$

and for the field

$$P = \mathrm{Tr}_a[\rho]. \tag{2.24}$$

Using the above initial conditions, atomic inversion $W(t)$ is given by

$$W(t) = \kappa_{aa} - \kappa_{bb} = \sum_{n=0}^{\infty} [\cos^2 2\theta_n |C_n|^2 + \sin^2 2\theta_n |C_n|^2 \cos(H_n t) \exp[-cH_n^2 t], \tag{2.25}$$

and population distribution in the field is given by

$$\begin{aligned} P_{nn}(t) = \tfrac{1}{2}[&|C_n|^2 + |C_{n-k}|^2 + \cos^2 2\theta_n |C_n|^2 - \cos^2 2\theta_{n-k} |C_{n-k}|^2 \\ &+ \sin^2 2\theta_n |C_n|^2 \cos(H_n t) \exp[-cH_n^2 t] - \sin^2 2\theta_{n-k} |C_{n-k}|^2 \\ &\times \cos(H_{n-k} t) \exp[-cH_{n-k}^2 t]], \end{aligned} \tag{2.26}$$

where H_n is given by

$$H_n = H_n^+ - H_n^-. \tag{2.27}$$

Off-diagonal elements of the atomic and field density matrix can also be obtained similarly. For example atomic coherence properties can be studied if we know the off-diagonal density matrix κ_{ab},

$$\begin{aligned} \kappa_{ab} = (1/2) \sum_{n=0}^{\infty} \{ &[\sin(\theta_{n+k} - \theta_n) + \sin(\theta_{nk} + \theta_n)] \rho_{1n\, n+k}(0) \exp[A^+_{n\, m+k} t] \\ &+ [\sin(\theta_{n+k} - \theta_n) - \sin(\theta_{nk} + \theta_n)] \rho_{4n\, n+k}(0) \exp[A^-_{n\, m+k} t] \\ &+ [\cos(\theta_{n+k} - \theta_n) + \cos(\theta_{nk} + \theta_n)] \rho_{2n\, n+k}(0) \exp[B^+_{n\, m+k} t] \\ &+ [-\cos(\theta_{n+k} - \theta_n) + \cos(\theta_{nk} + \theta_n)] \rho_{3n\, n+k}(0) \exp[B^-_{n\, m+k} t] \}, \end{aligned} \tag{2.28}$$

where

$$\begin{aligned} A^{\pm}_{nm} &= i(H_n^{\pm} - H_m^{\pm}) - c(H_n^{\pm} - H_m^{\pm})^2, \\ B^{\pm}_{nm} &= \mp i(H_n^{+} - H_m^{-}) - c(H_n^{+} - H_m^{-})^2. \end{aligned} \tag{2.29}$$

Field coherences can also be evaluated similarly as follows:

$$\begin{aligned} P_{nm}(t) = (1/2)[&\{[\cos(\theta_{m-k} - \theta_{n-k}) + \cos(\theta_{m-k} + \theta_{n-k})] \rho_{1n-k\, m-k}(t), \\ &+ [\cos(\theta_{m-k} - \theta_{n-k}) - \cos(\theta_{m-k} + \theta_{n-k})] \rho_{4n-k\, m-k}(t), \\ &+ [-\sin(\theta_{m-k} - \theta_{n-k}) - \sin(\theta_{m-k} + \theta_{n-k})] \rho_{2n-k\, m-k}(t), \\ &+ [-\sin(\theta_{m-k} - \theta_{n-k}) + \sin(\theta_{m-k} + \theta_{n-k})] \rho_{3n-k\, m-k}(t) \}, \\ &+ \text{similar terms with } m - k \to m \text{ and } n - k \to n]. \end{aligned} \tag{2.30}$$

From eq. (2.25) it is evident that the population inversion in the generalized JC model

decays to a steady state value

$$W(\infty) = \sum_{n=0}^{\infty} \cos^2 2\theta_n |C_n|^2. \tag{2.31}$$

Depending on the model, $\cos^2 2\theta_n$ assumes some finite value. For the ordinary JCM at exact resonance, $W(\infty) = 0$.

As the decoherence model is energy conserving, the loss of atomic energy is gained by the field and thus the average energy of the field is saturated to a higher value. Average photon number evolves as

$$\bar{n}(t) = \sum_{n=0}^{\infty} |C_n|^2 \{(n + k/2) - k/2 \cos^2 2\theta_n - k/2 \sin^2 2\theta_n \cos(H_n t) \exp[-cH_n^2 t]\}. \tag{2.32}$$

With the suppression of oscillation in the average photon number it reduces to a steady state value

$$\bar{n}(\infty) = \sum_{n=0}^{\infty} |C_n|^2 \{(n + k/2) - k/2 \cos^2 2\theta_n = \bar{n}(0) + (k/2)\left[1 - \sum_{n=0}^{\infty} |C_n|^2 \cos^2 2\theta_n\right]. \tag{2.33}$$

To investigate the sub-Poissonian nature of the field we need to know the average of the square of the photon number in the field which can be expressed as

$$\overline{n^2}(t) = \sum_{n=0}^{\infty} |C_n|^2 \{(n^2 + nk + k^2/2) - (n + k^2/2) \cos^2 2\theta_n - (n + k^2/2) \sin^2 2\theta_n \cos(H_n t) \exp[-cH_n^2 t]\}. \tag{2.34}$$

The sub-Poissonian characteristics of the field is exemplified for a particular Hamiltonian of atom-field interaction which is presented in § 3.

From eq. (2.28) it is evident that atomic coherence decays in time and at sufficiently large times it goes to zero. This is also true for the coherence terms of the field, i.e.,

$$P_{nm}(\infty) = 0, \quad \text{for } n \neq m. \tag{2.35}$$

Quadrature squeezing properties of the field can also be calculated by using eq. (2.30). The degree of squeezing can be measured by the squeezing parameter,

$$S_i = \frac{\langle (\Delta X_i)^2 \rangle - |\langle [X_1, X_2] \rangle|/2}{|\langle [X_1, X_2] \rangle|/2}, \tag{2.36}$$

where X_i can be defined as

$$X_1 = (1/2)(a e^{i\omega t} + a^\dagger e^{-i\omega t})$$
$$X_2 = (1/2i)(a e^{i\omega t} - a^\dagger e^{-i\omega t}). \tag{2.37}$$

A quadrature is said to be squeezed if $S_i < 0$. In terms of the expectation values $a, a^\dagger$ of $S_{1,2}$ can be defined as

$$S_1 = 2\langle a^\dagger a\rangle + 2\mathrm{Re}\langle a^2 e^{2i\omega t}\rangle - 4(\mathrm{Re}\langle a e^{i\omega t}\rangle)^2$$
$$S_2 = 2\langle a^\dagger a\rangle - 2\mathrm{Re}\langle a^2 e^{2i\omega t}\rangle - 4(\mathrm{Im}\langle a e^{i\omega t}\rangle)^2. \tag{2.38}$$

It can be shown easily that at the steady state,

$$S_1 = S_2 = 2\langle a^\dagger a\rangle > 0. \tag{2.39}$$

So at the steady state the generalized JC model shows there is no squeezing in any quadrature.

Finally, we would like to point out that decoherence on the generalized JC model acts as a relaxation on the dressed state coherence only. But it does not affect the dressed state populations unlike the case of ordinary spontaneous emission [10] where both the dressed state population and coherences decay in time and subsequently all the energy and coherence of the atom and the field drain out of the system at the steady state.

2b. Transient dynamics of atom or field for ordinary JCM

Here we explain the features which appear in the transient dynamics of atom or field due to this particular nature of decoherence. We take the ordinary JCM as the candidate to describe it. As the decoherence mechanism affects individual subsystems, e.g., the atom or the field in an almost similar manner we concentrate only on the dynamics of the atomic properties.

The Hamiltonian for the ordinary JCM describing a two-level atom interacting with a single mode field confined in a cavity within RWA is given by

$$H = \hbar\omega a^\dagger a + (\hbar\omega_0/2)\sigma_z + \hbar g(a^\dagger\sigma_- + a\sigma_+), \tag{2.40}$$

where $a(a^\dagger)$ is the annihilation (creation) operator of the single mode cavity field with frequency ω and σ_+, σ_- and σ_z are the usual spin half operators for the two-state atom with characteristic frequency ω_0. g is the coupling strength between the atom and the field mode.

To explain the transient dynamics the following expressions are in order for ordinary JCM. Calculating at exact resonance, i.e., when $\delta = \omega - \omega_0$ becomes zero and taking $\hbar = 1$ we obtain the atomic inversion $W(t)$ as

$$W(t) = \sum_{n=0}^{\infty} |C_n|^2 \cos[2g\sqrt{(n+1)}t]\exp[-4cg^2(n+1)t]. \tag{2.41}$$

The field population matrix elements evolve as

$$P_{nn}(t) = |C_n|^2\{\cos^2[g\sqrt{(n+1)}t]\exp[-4cg^2(n+1)t]$$
$$+(1/2)[1-\exp[-4cg^2(n+1)t]\} + |C_{n-1}|^2\{\cos^2[g\sqrt{nt}]\exp[-4cg^2nt]$$
$$+(1/2)[1-\exp[-4cg^2nt]\}. \tag{2.42}$$

The atomic and field coherence, i.e., off-diagonal atomic and field density matrix elements can be evaluated as

$$\kappa_{ab} = (C_n C^*_{n+1}/4) \sum_{n=0}^{\infty} \{\exp[A^+_{nn+1}t] - \exp[A^-_{nn+1}t] + \exp[B^+_{nn+1}t] - \exp[B^-_{n\,n+1}t]\}, \tag{2.43}$$

and

$$P_{nm} = (C_n C^*_m/4)\{\exp[A^+_{nm}t] + \exp[A^-_{nm}t] - \exp[B^+_{nm}t] + \exp[B^-_{nm}t]\} + \text{terms with } \{n \to n-1 \text{ and } m \to m-1\}, \tag{2.44}$$

where

$$A^{\pm}_{nm} = -i[\omega(n-m) \pm g\sqrt{(n+1)} \mp g\sqrt{(m+1)}] - c[\omega(n-m) \pm g\sqrt{(n+1)} \mp g\sqrt{(m+1)}]^2,$$

$$B^{\pm}_{nm} = \mp i[\omega(n-m) + g\sqrt{(n+1)} + g\sqrt{(m+1)}] - c[\omega(n-m) + g\sqrt{(n+1)} + g\sqrt{(m+1)}]^2.$$

An interesting aspect of the decoherence mechanism for ordinary JCM is that it affects the population and coherence properties of the individual subsystem in a different way. For the ordinary JCM, we have shown the collapse and revival dynamics of the atomic population for a parameter value of cg as 0.005, and photon number gets erased instantaneously. But the coherence properties of the atom or field decay at a slower rate. We show the time evolution of atomic population inversion and atomic coherence and the purity of the reduced atomic density in figures 1, 2 and 3 respectively. For numerical purpose, we assume $\omega/g = 2$. Note that $cg\hbar^2$ is a dimensionless quantity and here we assume $\hbar = 1$.

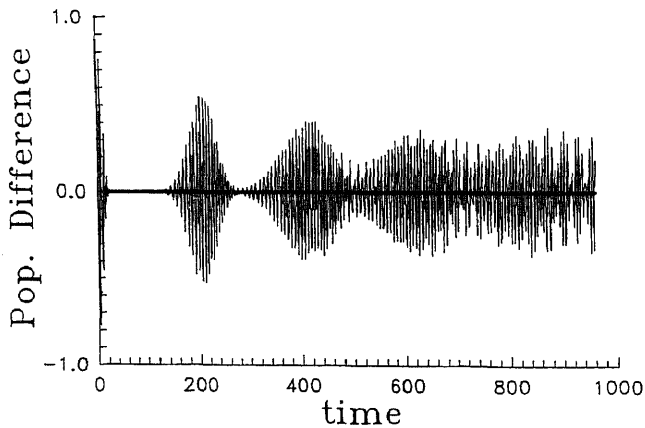


Figure 1. Population inversion $W(t)$ (eq. 2.41) is plotted with scaled time (gtx) when the field is initially in coherent state with $|x|^2 = 16$ (thinner line is pure JCM and thicker line is with decoherence ($cg\hbar^2 = 0.005$)).

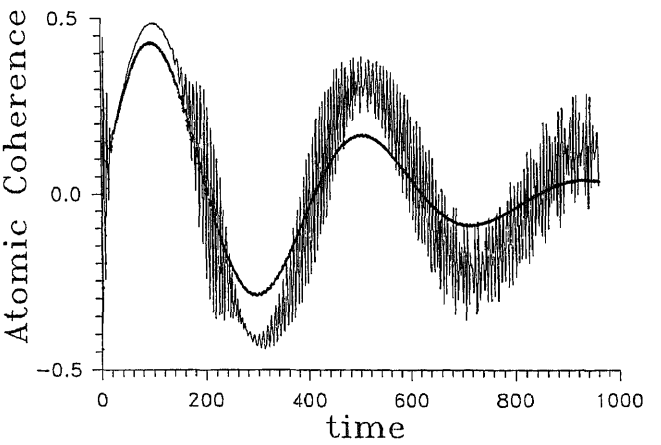


Figure 2. Atomic coherence κ_{ab} (eq. 2.43) is plotted with scaled time (gtx) when the field is initially in coherent state with $|x|^2 = 16$ (thinner line is pure JCM and thicker line with decoherence $cg\hbar^2 = 0.005$)).

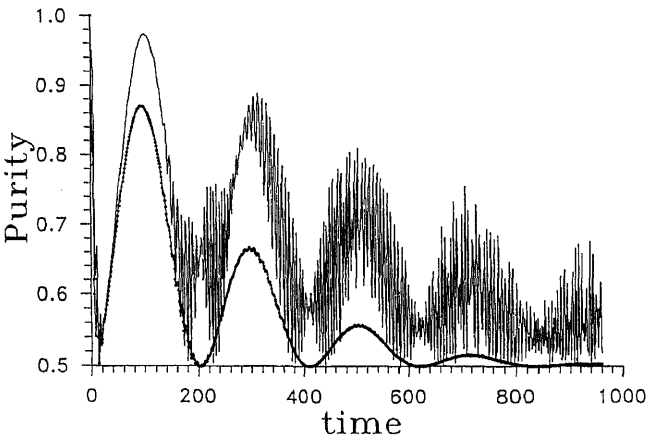


Figure 3. Purity of the atom Tr (ρ^2) is plotted with scaled time (gtx) when the field is initially in coherent state with $|x|^2 = 16$ (thinner line is pure JCM and thicker line is with decoherence $cg = 0.005$)).

In figure 1 we show the time evolution of $W(t)$ in eq. (2.41) with the value of $cg = 0.005$ and $|x|^2 = 16$. It is seen that the revival of atomic inversion gets erased and a permanent collapse of population appears. The oscillation in atomic inversion and photon number is well studied for the simple JCM system [1]. The oscillation appears due to the coherence built up during the atom-field interaction and thus the individual atom or field population depends on the coherence of the dressed state levels. Thus when a decoherence is introduced in the dressed state the oscillation disappears very quickly.

We show the atomic coherence evolution eq. (2.42) in figure 2 to illustrate that the effect of decoherence appears more slowly compared to that in the population terms. In the long time limit the atomic or the field coherences vanish.

Gea–Banacloche [13] has recently given a recipe for the preparation of atomic superposed state in the middle of the first collapsed region for ordinary JCM. To obtain the macroscopically superposed state of an atom one has to stay in the pure state i.e., $\mathrm{Tr}(\kappa^2) = 1$ which is the necessary and sufficient condition for a pure state. This is because unlike the state vector, the density operator does not really describe an individual system but rather an ensemble of identically prepared atoms.

Keeping in view of Gea–Banacloche's observations we study how the decoherence affects the purity of the atom. This is displayed in figure 3. Here it is quite clear that although the atom is almost permanently collapsed for a low value of the decoherence parameter c, the purity of the atom oscillates between 1 (pure state) and 0.5 (maximally mixed ensemble). From figure (3) it is apparent that the purity does not get affected much although the collapse occurs in the inversion very rapidly. In the first half-collapse-time of JCM, the purity is not far from one; also it depends on the average initial photon number. In the long time the purity gets decayed sequentially after many Rabi cycles. The only effect which can be observed due to decoherence is that the profile of time evolution of purity gets smoothened over large oscillations due to the suppression of dressed state coherences. Thus we can conclude that the superposed state of the atom in a JCM system in the middle of the first collapse time is robust even under the influence of environmental pure decoherence on the atom-field composite system as long as the rate of decoherence is not large.

When the JCM system undergoes decoherence, two distinct interference effects arise. The first one is the ordinary interference of dressed state coherences in pure JCM which is dynamical in nature and is reversible, thereby giving recurrence of collapse and revival of the population and a transient oscillation in the coherence properties of atom and field over a slowly varying profile. The second, interference effect is due to the decay of coherence of dressed states which is irreversible in character. The mechanism of decoherence as introduced here affects the population and the coherence in different ways. From eqs (2.41), (2.42) and eqs (2.43), (2.44) it is clear that the decay rate in population and coherence of an individual subsystem depends on the time scale of oscillation in the free evolution of the dressed state of the JCM system. In the free evolution of population there is a fast oscillation of collapse and revival but in the coherence terms of an individual subsystem there is a slower time scale of oscillation along with a fast oscillation. Thus in the decay rate of coherence these two components will arise unlike the case of population where only one kind of decay rate will appear which is fast. So when the decoherence mechanism is introduced in JCM, it first kills the fast oscillation in population and coherence properties of the atom and field. This gives steady redistributed energy values of atom and field and a smooth profile of slow evolution of coherence terms. Then the decoherence mechanism acts slowly to wash out the coherence of individual subsystems and takes a few Rabi cycles to reach to their vanishing values in the steady state. The diminishing curve of atomic purity in figure 3 can be easily understood in terms of the number of Rabi cycles needed for the complete wash out of atomic coherence for a fixed value of the parameter c.

interaction of atom and field with Kerr nonlinearity in the cavity. The Hamiltonian is given by

$$H = \hbar\omega a^\dagger a + \hbar q a^{\dagger^2} a^2 + (\hbar\omega_0/2)\sigma_z + \hbar g(a^{\dagger k}\sigma_- + a^k\sigma_+). \tag{3.1}$$

Here q is the nonlinearity parameter for the Kerr medium and the two-level atom is interacting with the field via a k-photon mediated process. For $k = 1$ and $q = 0$ this model reduces to the ordinary JCM in eq. (2.40).

With the help of §2(a) the various parameters can be identified as follows:

$$b_n = -\omega_0/2 + n\omega + qn(n-1), a_n = \omega_0/2 + n\omega + qn(n-1),$$
$$\Delta_n = \delta - q(2nk + k^2 - k), \quad \delta = \omega_0 - k\omega,$$
$$R_n = g\sqrt{\frac{(n+k)!}{n!}},$$
$$H_n^{\pm} = (n + k/2)\omega + q[n^2 - n + nk + k^2/2 - k/2] \pm \{(\Delta_n/2)^2 + R_n^2\}^{1/2}$$

and

$$\tan\theta_n = \frac{g[(n+k)!/n!]^{1/2}}{(\delta/2+q(nk+k^2/2-k/2))+\{g^2(n+k)!/n! + [\delta/2 + q(nk + k^2/2 - k/2)]^2\}^2}. \tag{3.2}$$

For $k = 1$, this Hamiltonian is studied in detail in the context of micromaser [14], population trapping [15] and it shows beating [16] in population inversion instead of conventional collapse and revival due to its intensity dependent detuning [17]. For $\delta = 0$ and $q = g/2$ it shows an interesting behaviour [15] in this model due to the fact that the Rabi frequency in this case becomes $g(n/2 + 1)^{1/2}$ instead of $g(n + 1)^{1/2}$ in conventional JCM.

Due to the intensity dependent detuning, here $\tan\theta_n$ is not equal to 1 even if $\delta = 0$. So the population inversion in the steady state in this case can be evaluated as

$$W(\infty) = \sum_{n=0}^{\infty} \cos^2 2\theta_n |C_n|^2. \tag{3.3}$$

So depending on the nonlinearity parameters q and k, atomic inversion can be saturated to a value higher than 0.5. Thus in the presence of nonlinearity, the transfer of energy from atom to field is less than that in the normal JCM.

The Mandel Q parameter which is defined as

$$Q = \frac{\overline{\Delta n^2} - \bar{n}}{\bar{n}} \tag{3.4}$$

can be obtained from the steady state values of $\bar{n}(\infty)$ and $\overline{n^2}(\infty)$ as

$$\bar{n}(\infty) = \sum_{n=0}^{\infty} |C_n|^2 \{(n + k/2) - k/2 \cos^2 2\theta_n\}, \tag{3.5}$$

$$\overline{n^2}(\infty) = \sum_{n=0}^{\infty} |C_n|^2 \{(n^2 + nk + k^2/2) - (n + k^2/2) \cos^2 2\theta_n\}. \tag{3.6}$$

So even if $q = 0$ and $\delta = 0$, $\cos 2\theta_n$ is equal to zero and thus

$$Q(\infty) = -\frac{1}{4\bar{n}(0) + k/2}. \tag{3.7}$$

If $q/g = 0.5, \delta = 0$ and $k = 1$, in this case also $\cos 2\theta_n = 0$ and the value of $Q(\infty)$ is the same as above.

Now if $\tan\theta_n < 1$, for $q \neq 0, \delta \neq 0, k \neq 1$, the steady state Q-parameter is always less than 0, which means that the sub-Poissonian nature is not erased.

4. More generalized JC models

Here we present a more generalized JCM which not only accounts for the various ordinary JCM of atom-field interactions as described in the earlier sections but also the models describing a class of spin-orbit interactions.

Here we write the total Hamiltonian of the spin-orbit system as

$$H = \sum_{n=n_0}^{n_m} b_n |n\,b\rangle\langle n\,b| + a_n |n\,a\rangle\langle n\,a| + \sum_{n=n_0}^{n_m-1} R_n |n\,a\rangle\langle n+1b| + R_n^* |n+1b\rangle\langle n\,a|, \tag{4.1}$$

where $|a\rangle(|b\rangle)$ represents the upper (lower) spin state and n is the label for the angular momentum states. Unlike the ordinary JCM states $|n\rangle$ where n runs from 0 to ∞, for the angular momentum states n runs from $-l$ to $+l$ with l the orbital angular momentum. In this case $n_0 = -l$ and $n_m = +l$. For simplicity, we assume one quantum of energy is transferred at a time from spin to orbit states or vice versa. R_n is the coupling constant depending on the algebraic property of the operators, $\sum_{n=n_0}^{n_m-1} |n\rangle\langle n+1|$, $\sum_{n=n_0}^{n_m-1} |n+1\rangle\langle n|$ and $\sum_{n=n_0}^{n_m} |n\rangle\langle n|$.

Unlike the ordinary JCM, this Hamiltonian has two zero point energy states which do not couple with other states. These states are $|n_0\,b\rangle$ and $|n_m\,a\rangle$. These two states are diagonal in H as

$$\begin{aligned} H|n_0\,b\rangle &= b_{n_0}|n_0\,b\rangle, \\ H|n_m\,a\rangle &= a_{n_m}|n_m\,a\rangle. \end{aligned} \tag{4.2}$$

For all other states with $n \neq n_0, n_m$, H can be diagonalized as

$$H|\psi_{n_1}^{\pm}\rangle = H_{n_1}^{\pm}|\psi_{n_1}^{\pm}\rangle, \tag{4.3}$$

where

$$H_{n_1}^{\pm} = \frac{b_{n_1+1} + a_{n_1}}{2} \pm \Omega_{n_1} \tag{4.4}$$

and

$$\Omega_{n_1}^2 = (\Delta_{n_1}/2)^2 + |R_{n_1}|^2; \tag{4.5}$$

with detuning

$$\Delta_{n_1} = a_{n_1} - b_{n_1+1}. \tag{4.6}$$

The eigenstates $|\psi_n^{\pm}\rangle$ are given by

$$\begin{aligned} |\psi_n^+\rangle &\equiv |+n_1\rangle = \cos\theta_{n_1}|b\,n_1+1\rangle + \sin\theta_{n_1}|a\,n_1\rangle, \\ |\psi_n^-\rangle &\equiv |-n_1\rangle = -\sin\theta_{n_1}|b\,n_1+1\rangle + \cos\theta_{n_1}|a\,n_1\rangle, \end{aligned} \tag{4.7}$$

with θ_{n_1} obtained from

$$\tan\theta_{n_1} = \frac{R_{n_1}}{\Delta_{n_1}/2 + \Omega_{n_1}}. \tag{4.8}$$

With these discussions in mind one can generalize the above Hamiltonian H for multistate spin systems instead of two state spin system. A more general Hamiltonian in that case can be written as

$$H = \sum_{s=s_0}^{s_m}\sum_{n=n_0}^{n_m} a_{n\,s}|n\,s\rangle\langle n\,s| + \sum_{s=s_0}^{s_m-k}\sum_{n=n_0}^{n_m-k} R_{n\,s}|n\,s+k\rangle\langle n+k\,s| + R_{n\,s}^*|n+k\,s\rangle\langle n\,s+k|, \tag{4.9}$$

where k quantum is exchanged between the spin-orbit interacting system and this Hamiltonian is also diagonalized in a similar way. Here we will not dwell on further applications of this Hamiltonian (4.9).

5. A hydrogen atom in presence of a magnetic field undergoing pure decoherence—effect on spin dynamics

Here we present a special case of spin-orbit interaction discussed in the previous section – a model of hydrogen atom in presence of a constant magnetic field [18, 19]. The Hamiltonian can be written as

$$H = H_0 + H_1, \tag{5.1}$$

where

$$H_0 = p^2/2m + V(r) \tag{5.2}$$

is the usual free part of hydrogen atom with $V(r)$ the central potential and p and m are the momentum and mass of the electron. The interaction term H_1 can be written as

$$\begin{aligned} H_1 &= \beta(L_z + \sigma_z) + \alpha\, L.S \\ &= \beta L_z + [(\alpha/2)L_z + \beta]\sigma_z + (\alpha/2)(L_-\sigma_+ + L_+\sigma_-), \end{aligned} \tag{5.3}$$

where β is the strength of the magnetic field applied in the z-direction and (L_z, L_+, L_-) and $(\sigma_z, \sigma_+, \sigma_-)$ are the usual orbital and spin angular momentum operators respectively. α is the spin-orbit coupling constant which depends on the central potential $V(r)$. We solve here only H_1 for a fixed value of l.

From the previous section one can immediately identify that

$$b_{n_l} = -\beta + \beta n_l - \alpha n_l/4, \tag{5.4}$$

and

$$a_{n_l} = \beta + \beta n_l + \alpha n_l/4. \tag{5.5}$$

Δ_{n_l} is defined as

$$\Delta_{n_l} = a_{n_l} - b_{n_l} = \beta + \alpha(2n_l + 1)/4 \tag{5.6}$$

and

$$H^{\pm}_{n_l} = \frac{\beta(2n_l+1)}{2} - \frac{\alpha}{8} \pm \frac{\alpha}{2}\sqrt{[(l-n_l)(l+n_l+1)]}, \tag{5.7}$$

where l is the orbital angular momentum quantum number, and θ_{n_l} is given by

$$\tan\theta_{n_l} = \frac{(\alpha/2)\sqrt{[(l-n_l)(l+n_l+1)]}}{\Delta_{n_l}/2 + \Omega_{n_l}}. \tag{5.8}$$

In this case $n_0 = -l$ and $n_{m_l} = +l$.

Using the above results one can obtain exactly the energy levels of any state of hydrogen atom in presence of a magnetic field by avoiding the approximations which are commonly made. One such approximation is the Zeeman approximation, where the magnetic field is assumed to be so weak that the terms with magnetic field can be treated as perturbation. The other approximation is the Paschen–Back limit when the magnetic field is so strong that the spin-orbit interaction is taken as perturbation.

Now we consider the Hamiltonian of a hydrogen atom in a magnetic field as the system Hamiltonian and assume a model of decoherence as in § 2(a) for this particular system. We discuss the dynamics of spin in presence of magnetic field and investigate the effect of pure decoherence on the spin motion of hydrogen atom in a magnetic field.

Using the same notations as used in § 2(a), the Liouville equation (2.11) for this Hamiltonian can be solved to obtain the spin inversion $W(t)$ as

$$\begin{aligned} W(t) =& (1/2)[(\rho_{1n_mn_m}(t) + \rho_{4n_mn_m}(t)) + \cos 2\theta_{n_m}(\rho_{1n_mn_m}(t) - \rho_{4n_mn_m}(t)) \\ & - \sin 2\theta_{n_m}(\rho_{2n_mn_m}(t) + \rho_{3n_mn_m}(t))] - \sum_{n_l=n_0}^{n_m} \cos 2\theta_{n_l}[(\rho_{1n_ln_l}(t) \\ & - \rho_{4n_ln_l}(t) - \sin 2\theta_{n_l}(\rho_{2n_ln_l}(t) + \rho_{3n_ln_l}(t))]; \end{aligned} \tag{5.9}$$

this expression is valid for any l.

We take the orbital angular momentum value $l = 1$, so the magnetic quantum number n_l can assume three values $-1, 0, +1$. The initial condition of the spin-orbit system is taken as

$$\psi(0) = (1/\sqrt{6})(|n_{+1}\rangle + |n_0\rangle + |n_{-1}\rangle)(|a\rangle + |b\rangle). \tag{5.10}$$

Now for $l = 1$, the initial condition can be rewritten in terms of the dressed state densities as follows:

$$\begin{aligned}
\rho_{1n_l n_l} &= (1 + \sin 2\theta_{n_l})/6, \\
\rho_{4n_l n_l} &= (1 - \sin 2\theta_{n_l})/6, \\
\rho_{2n_l n_l} &= (\cos 2\theta_{n_l})/6, \\
\rho_{3n_l n_l} &= (\cos 2\theta_{n_l})/6.
\end{aligned} \tag{5.11}$$

Finally $W(t)$ can be expressed as

$$\begin{aligned}
W(t) = &\left[1/6 + (1/12)\sin 4\theta_{n_m} - (1/6)\sum_{n_l=n_0}^{n_m} \sin 4\theta_{n_l}\right] \\
&- (1/12)\sin 4\theta_{n_m} \exp(-cH_{n_m}^2 t)\cos(H_{n_m} t) \\
&+ \sum_{n_l=n_0}^{n_m} (1/12)\sin 4\theta_{n_l} \exp(-cH_{n_l}^2 t)\cos(H_{n_l} t),
\end{aligned} \tag{5.12}$$

where H_{n_l} is defined as

$$H_{n_l} = H_{n_l}^+ - H_{n_l}^-. \tag{5.13}$$

So for $c = 0$, that is when there is no decoherence, $W(t)$ has an oscillatory pattern composed of interference of $\cos(H_{n_l} t)$ terms and may show a collapse and revival type behavior as in atom-field JCM case. But when decoherence is present, i.e., $c \neq 0$, the spin inversion relaxes with time. At the steady state the value of $W(\infty)$ depends on the strength of the magnetic field and spin-orbit coupling strength which is apparent from eq. (5.12).

6. Summary and conclusion

We have investigated the effect of environment induced pure decoherence on the generalized Jaynes–Cummings model in the atom-field interaction case. It is shown that at the steady state, transient oscillations in atom and field observables are suppressed but all the nonclassical properties are not erased by the decoherence. For example, it is seen that an initial Poissonian field acquires a nonclassical character in the steady state. Steady state observables are shown to be model dependent. This is shown with an example of a multiphoton JCM in presence of a Kerr nonlinearity in the cavity.

It is also shown that pure decoherence affects the population and coherence properties of the individual subsystem in a different way. For ordinary JCM we have shown that for a small value of the parameter c, atomic population inversion collapses very fast but the purity and other phase related quantities of atom take a few Rabi cycles to reach their steady state. Thus the atomic superposition in the middle of the first-half-collapse time is robust even under the influence of environment induced pure decoherence.

We have studied the effect of decoherence on more generalized JC models which not only account for the atom-field interactions but also a class of spin-orbit interactions in

the framework of JCM. As a particular example, the dynamics of spin is studied for the hydrogen atom in presence of a magnetic field undergoing decoherence. A more generalized JCM Hamiltonian and its solution are also proposed which can account for multiple spin states.

It is interesting to note that the pure decoherence can not erase all nonclassical features of atom and field unlike the case of spontaneous emission, where along with the decoherence, an energy dissipation guarantees the steady state with no nonclassical behaviour in either subsystem.

Nevertheless, the structure of the Liouville equation (2.11) has some general importance because various mechanisms of stochastic interactions with the system in the Born–Markov limit lead to this kind of equation. We hope that the generalized JCM undergoing decoherence may be worthwhile to investigate whenever a stochastic modulation of energy of the system is present.

Acknowledgement

Partial financial support from National Science Council of R.O.C. is thankfully acknowledged. We are grateful to the unknown referee for useful comments.

References

[1] See for example a recent review by, B W Shore and P L Knight, *J. Mod. Opt.* **40**, 1195 (1993) and references therein
[2] E A Hinds, in *Advances in atomic and molecular and optical physics* edited by D Bates and B Bederson (NY, Academic) **28**, p. 237
S Haroche and D Kleppner, *Phys. Today*, **42**, 24 (1989)
[3] A Tameshtit and J E Sipe, *Phys. Rev.* **A45**, 8280 (1992); **47**, 1697 (1993)
[4] J A Wheeler and W H Zurek, *Quantum theory of measurement.* (Princeton, 1986)
A Venugopalan, D Kumar and R Ghosh, *Physica* **A220**, 563 (1995)
[5] G J Milburn, *Phys. Rev.* **A44**, 5401 (1991); **47**, 2415 (1993)
[6] A Venugopalan, D Kumar and R Ghosh, *Physica* **A220**, 576 (1995)
[7] See, for example, a recent review on sub-Poissonian light by L Davidovich, *Rev. Mod. Phys.* **68**, 127 (1996) and references therein
[8] G J Milburn and D F Walls, *Phys. Rev.* **A28**, 2646 (1983)
[9] R Alicki. *Phys. Lett.* **A103**, 247 (1984)
[10] W H Louisell, *Quantum Statical properties of radiation* (NY, John Wiley, 1973)
G Gangopadhyay and D S Ray, in *Advances in multiphoton processes and spectroscopy* edited by S H Lin, A A Villaeys and Y Fujimura **8**, (1993)
[11] H Moya-Cessa, V Buzek, M S Kim and P L Knight, *Phys. Rev.* **A48**, 3900 (1993)
X Chen and L M Kuang, *Phys. Lett.* **A191**, 18 (1994)
[12] C Cohen-Tannoudji, J Dupont-Roc and G Grynberg, *Atom-Photon Interaction* (NY, Wiley, 1992)
[13] J Gea-Banacloche, *Phys. Rev. Lett.* **65**, 3385 (1990)
[14] B Deb and D S Ray, *Phys. Rev.* **A48**, 3191 (1993)
[15] A Bandyopadhyay and G Gangopadhyay, *J. Mod. Opt.* **43**, 487 (1996)
[16] S Du and C Gong, *Phys. Rev.* **A17**, 779 (1994)
[17] V Buzek and I Jex, *Opt. Commun.* **78**, 425 (1990)
[18] J J Sakurai, *Modern quantum mechanics* (Addison-Wesley, Reading, MA, 1985)
[19] S Yu, H Rauch and Y Zhang, *Phys. Rev.* **A52**, 2585 (1995)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 4,
October 1997
pp. 417–420

Isothermal fluid sphere: Uniqueness and conformal mapping

NARESH DADHICH
Inter-University Centre for Astronomy and Astrophysics, P.O. Box 4, Pune 411 007, India
e-mail: nkd@iucaa.ernet.in

MS received 15 April 1997

Abstract. We prove the theorem: A necessary and sufficient condition for a spacetime to represent an isothermal fluid sphere (linear equation of state with density falling off as inverse square of the curvature radius) without boundary is that it is conformal to a spacetime of zero gravitational mass ('minimally' curved).

Keywords. Fluid spheres; isothermal fluid; conformal mapping.

PACS Nos 04.20; 98.80

Except for Brinkman's theorem [1] conformally relating two Einstein spaces, only spacetimes conformal to flat spacetime have been considered. The well-known examples of such spacetimes, that are also of great astrophysical and cosmological interest, are the Friedman–Robertson–Walker (FRW) flat model of the Universe and the Schwarzschild solution describing interior of a star in hydrostatic equilibrium. These spacetimes have distinguishing physical properties, like isotropy and homogeneity for the former and uniform density for the latter. In this note we shall establish a unique association between isothermality of fluid sphere with conformal mapping of spacetime. Such a clear characterization of physical behaviour with geometric property is rather very rare.

Here the base spacetime is of course not flat but can be thought as 'minimally' curved [2]. It is a spherically symmetric spacetime from which radial acceleration as well as tidal acceleration for radial motion have been anulled out but curvature is not zero, which manifests only in tidal acceleration for transverse motion. Though it is not a solution of the Einstein vacuum equation, it presents an interesting physical situation with the property, $R_{ik}u^iu^k = 0, u^iu_i = 1$. This means spacetime is free of active gravitational mass. Further it has all but one curvature zero, which is also an invariant for spherical symmetry, and it is at any given radius proportional to curvature of sphere of that radius [3]. This is why we have termed it as 'minimally' curved spacetime (MCS) [2, 4]. We have very recently considered metric in the Kerr–Schild (KS) form with a view to find perfect fluid solutions [5]. It is however well-known that perfect fluid is not compatible with the KS form and hence Senovilla and Sopuerta [6], and Martin and Senovilla [7] have generalized it by replacing flat metric by conformally flat. This would mean that

A metric in KS form is given by

$$g_{ij} = \eta_{ij} + 2Hl_i l_j, \tag{1}$$

where H is a scalar field and l_i is a null vector relative to both g_{ij} and η_{ij}. For spherically symmetric vacuum solution; i.e. the Schwarzschild solution, H satisfies the Laplace equation $\nabla^2 H = 0$. It is rather interesting to note that spacetime is not flat unless $H = 0$; i.e. $H = \text{const.} \neq 0$, represents a curved spacetime and this is the MCS discussed above.

Let us consider a metric conformal to MCS by taking H constant. Recall that for perfect fluid seed metric has to be different from vacuum. So we take H constant and write

$$\bar{g}_{ij} = e^{2U}(\eta_{ij} + 2Hl_i l_j), H = \text{const.} \tag{2}$$

which for spherical symmetry can be brought to the orthogonal form to read [5]

$$\mathrm{d}s^2 = e^{2U}(-\mathrm{d}t^2 + k^2\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2). \tag{3}$$

Here $U = U(r,t)$ and $k^2 = (1+2H)^2/(1-2H)$ and the base metric is MCS.

First, the perfect fluid conditions in the comoving coordinates ($u_i = e^U\delta_i^0$) imply $U = U(r)$, and they yield the general solution for unbounded distribution [5],

$$e^U = r^{-n} \tag{4}$$

with

$$8\pi\rho = \frac{n(n-2)}{k^2 r^{2(1-n)}}, \quad \rho\frac{n-2}{n}p, \tag{5}$$

where $k^2 = 1 + 2n(n-2)$ and removable constants have been transformed away. This is the general solution which represents isothermal fluid as it admits a linear equation of state and density falls off as inverse square of the curvature radius, $R = r^{1-n}$. It is important to note that the metric (3) admits this unique solution (4) representing isothermal fluid sphere without boundary [5].

Thus the metric ansatz (3) with the unique solution (4) is the sufficient condition for isothermal fluid sphere without boundary. We shall now turn to the necessary condition. For that we have to prove that all unbounded isothermal fluid sphere solutions can always be cast in the metric ansatz (3). Let us consider the general spherically symmetric metric

$$\mathrm{d}s^2 = -e^{\nu}\mathrm{d}t^2 + e^{\lambda}\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2), \tag{6}$$

where λ and ν are functions of r only. With this we have for perfect fluid in comoving coordinates ($u_i = e^{\nu/2}\delta_i^0$),

$$8\pi\rho = \frac{1}{r^2}[1 + e^{-\lambda}(r\lambda' - 1)], \tag{7}$$

$$8\pi p = \frac{1}{r^2}[-1 + e^{-\lambda}(r\nu' + 1)], \tag{8}$$

$$2\nu'' + \nu'^2 - \lambda'\nu' - \frac{2}{r}(\nu' + \lambda') + \frac{4}{r^2}(e^{\lambda} - 1) = 0, \tag{9}$$

where $\lambda' = \partial\lambda/\partial r$. Then we obtain the general solution for a regular e^{λ} at $r = 0$ [8],

$$e^{\lambda} = \text{const.} = k_1^2, \quad e^{\nu} = r^{-2m}. \tag{10}$$

This belongs to the Tolman class of solutions [9]. Thus we have obtained the general solution for unbounded isothermal fluid sphere and the metric (6) will read

$$ds^2 = r^{-2m}[-dt^2 + k_1^2 r^{2m} dr^2 + r^{2(1+m)}(d\theta^2 + \sin^2\theta d\varphi^2)]. \tag{11}$$

By redefining the radial coordinate as $\bar{r} = r^{1+m}$, the above metric on dropping overhead bar takes the form

$$ds^2 = r^{-2m/(1+m)}[-dt^2 + k_1^2 dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)] \tag{12}$$

which is exactly in the required ansatz form (3).

We have thus proved the theorem: A necessary and sufficient condition for an isothermal fluid sphere without boundary is that its spacetime metric is conformal to the minimally curved spacetime, characterized by vanishing of gravitational mass ($R_{ik}u^i u^k = 0$),

$$ds^2 = dt^2 + k^2 dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2), k = \text{const.}. \tag{13}$$

It is quite remarkable that there exists a one-to-one association between isothermality of fluid sphere and conformally MCS. Isothermality of fluid sphere picks up uniquely the geometric property, conformally MCS. This is really very interesting because it is a rare case of a specific physical property singling out a geometric property. Like the isothermal case, stresses generated by k in MCS also fall off as $T_0^0 = T_1^1 \sim r^{-2}$. These stresses have been identified with gravitational global monopole [10]. Global monopole is an exotic object which is supposed to be created when global $O(3)$ symmetry is spontaneously broken into $U(1)$ in phase transitions in the very early Universe. The simplest way to get to MCS is that it is a spacetime corresponding to a constant potential [4]. Recall that H in (1) satisfies the Laplace equation and it corresponds to the Newtonian potential. However we confess that its interpretation as potential may not be generally agreed upon.

Viewing MCS as a small departure (as it differs from flat spacetime only in tidal acceleration for transverse motion) from flat spacetime, it is interesting to note that for its conformal spacetime it does not permit non-static perfect fluid solution. It admits the general solution having three free parameters, which can represent bounded fluid spheres as well as unbounded isothermal fluid spheres [5]. In contrast to conformally flat metric, conformally MCS is highly constrained. Even though MCS may be a small deviation from flat spacetime, their conformal spacetimes have very different physical properties.

Isothermal fluid structures have been considered in astrophysics for a long time as an equilibrium approximation to more complicated systems approaching dynamical relaxed state. Recently the spacetime (11) has been argued as an ultimate end state of the Einstein–deSitter universe [8]. It is envisioned that the Einstein–deSitter model asymptotically tends to an expansion free state, and then it condenses into a stable isothermal fluid sphere.

References

[1] H W Brinkman, *Math. Ann.* **91**, 269 (1924)
[2] N Dadhich and K Narayan, *An ansatz for spacetimes of zero gravitational mass: Global monopoles and textures* (to be submitted)
[3] N Dadhich, An analysis of some unsolved problems in general relativity, Ph.D. thesis (Pune University, 1970) (unpublished)
[4] N Dadhich, *On understanding the Schwarzschild field* (to be submitted)
[5] N Dadhich, *Gen. Relativ. Gravit.* **28**, 1455 (1996)
[6] J M M Senovilla and C F Sopuerta, *Class. Quantum Gravit.* **11**, 2073 (1994)
[7] J Martin and J M M Senovilla, *J. Math. Phys.* **27**, 265 (1986)
[8] W C Saslaw, S D Maharaj and N Dadhich, *Astrophys. J.* **471**, 571 (1996)
[9] R C Tolman, *Phys. Rev.* **55**, 365 (1934)
[10] M Barriola and A Vilenkin, *Phys. Rev. Lett.* **63**, 341 (1989)

Investigation of the boundary value problem corresponding to a generalized Skyrme lagrangian

RAJESH KUMAR and P K CHATTOPADHYAY
Department of Physics, Maharshi Dayanand University, Rohtak, Haryana 124 001, India

MS received 14 August 1996; revised 18 February 1997

Abstract. A detailed numerical analysis of the boundary value problem resulting from the most general Skyrme type lagrangian containing up to quartic terms in field gradients is presented. The additional parameters in the lagrangian can be related to pion-pion scattering lengths. It is found that solutions to the boundary value problem does not exist for all values of the parameters and in particular, for the values predicted from pion-pion scattering data. Physical quantities of the nucleon are calculated for the highest possible values of the parameters admitting a solution and are compared with the corresponding values for the Skyrme model and experimental values.

Keywords. Skyrmion; Skyrme lagrangian; chiral symmetry; heavier mesons.

PACS No. 12.35

1. Introduction

The Skyrme model [1, 2] has emerged as an attractive model of baryons. In this model, baryons are solitonic solutions of non-linear pion fields. Since the original work of Skyrme and the work of Adkins *et al* [3] who calculated the physical quantities of the nucleon from the model, various generalizations have been attempted. On the one hand there are theories involving mesons heavier than the pion [4–8], on the other, terms higher order in field gradients have been included [9, 10]. The intimate connection between the higher order terms and heavier mesons has also been discussed [11, 12]. Weinberg [13] has written down the most general Skyrme type lagrangian involving quartic terms in the field gradients. In this paper we report results of a systematic investigation of the boundary value problem arising out of the Weinberg lagrangian. We also throw light on the inter-relationship of various Skyrme type lagrangians. Comparison of various calculated physical quantities with experimental results show that though there is not much change in the fitted value of the parameter F_π (the pion decay constant), there is some improvement over the Skyrme value in the case of other physical quantities.

In § 2 the inter-relationship of various Skyrme lagrangians is discussed. The boundary value problem is set up in § 3 and the solution of the boundary value problem is discussed in § 4. This is followed by § 5 which contains the calculated values of the physical quantities of the nucleon and their comparisons with the Skyrme and experimental values. § 6 contains the concluding remarks.

2. Generalized Skyrme lagrangians

The Skyrme lagrangian density is given by

$$\mathcal{L}_{\rm sk} = \frac{1}{16} F_\pi^2 \,\mathrm{Tr}(L_\mu L^{\mu+}) + \frac{1}{32e^2}\,\mathrm{Tr}\,[L_\mu, L_\nu]_-^2, \tag{1}$$

where $L_\mu = u^+\partial_\mu u, u$ being a $SU(2)$ matrix whose elements are meson fields and F_π is the pion decay constant. Donoghue, Golowich and Holstein [14] (DGH) modified this by adding an additional term quartic in fields gradients;

$$\mathcal{L}_{\rm DGH} = \mathcal{L}_{\rm sk} + \frac{\gamma}{8e^2}\,[\mathrm{Tr}\,(\partial_\mu L^\mu)]^2. \tag{2}$$

The parameter γ can be related to pion-pion scattering data.

Weinberg [13] has proposed that the most general Skyrme type lagrangian density involving up to quartic terms in the field gradients, under very general restrictions of locality, Lorentz invariance, chiral symmetry and CPT theorem is given by

$$\mathcal{L} = \mathcal{L}_{\rm DGH} + \frac{\beta}{32e^2}\,\mathrm{Tr}[L_\mu, L_\nu]_+^2. \tag{3}$$

Other combinations of L_μ, L_ν up to quartic terms in field gradients can be reduced to the form (3) by using the Maurer–Cartan identity

$$\partial_\mu L_\nu - \partial_\nu L_\mu + [L_\mu, L_\nu]_- = 0. \tag{4}$$

With $\beta = 0$ and $\gamma = 0$, $\mathcal{L}$ reduces to the Skyrme lagrangian density and one obtains $\mathcal{L}_{\rm DGH}$ from $\mathcal{L}$ by putting $\beta = 0$ and $\gamma \neq 0$.

The first term in (1) constitute the non-linear sigma model lagrangian. The other terms in (3) are quartic-derivative corrections to the model. An important observation by Gasser and Leutwyler [15] is that in the chiral limit, there exist only two independent quartic-derivative correction terms to the non-linear sigma model lagrangian viz. $[\mathrm{Tr}(\partial_\mu u \partial^\mu u^+)]^2$ and $[\mathrm{Tr}(\partial_\mu u \partial^\nu u^+)]^2$.

It is therefore necessary to analyse lagrangians (1), (2) and (3) in the light of this observation. $\mathcal{L}$ in (3) can be expressed in terms of u and u^+ using the identity $u^+\partial_\mu u = -\partial_\mu u^+ u$. Thus

$$\begin{aligned}\mathrm{Tr}(L_\mu L_\nu L^\nu L^\mu) &= \mathrm{Tr}(\partial^\mu u \partial_\mu u^+ \partial_\nu u \partial^\nu u^+) \\ &= \tfrac{1}{2}[\mathrm{Tr}(\partial^\mu u \partial_\mu u^+)]^2, \end{aligned} \tag{5}$$

$$\begin{aligned}\mathrm{Tr}(L_\mu L_\nu L^\mu L^\nu) &= \mathrm{Tr}(\partial_\mu u^+ \partial_\nu u \partial^\mu u^+ \partial^\nu u) \\ &= -\tfrac{1}{2}[\mathrm{Tr}(\partial_\mu u^+ \partial^\mu u)]^2 + [\mathrm{Tr}(\partial_\mu u^+ \partial^\nu u)]^2 \end{aligned} \tag{6}$$

and

$$[\mathrm{Tr}(\partial_\mu L^\mu)]^2 = [\mathrm{Tr}(\partial_\mu u^+ \partial^\mu u)]^2, \tag{7}$$

since in the chiral limit $\partial_\mu \partial^\mu u = 0$. The last steps in (5) and (6) are the consequences of $SU(2)$ trace identities [16]. Using (5), (6) and (7), $\mathcal{L}$ in eq. (3) can be written as

$$\mathcal{L} = \frac{1}{16} F_\pi^2 \operatorname{Tr}(\partial_\mu u \partial^\mu u^+) - \frac{1}{16e^2}(1-2\gamma)[\operatorname{Tr}(\partial_\mu u^+ \partial^\mu u)]^2$$
$$+ \frac{1}{16e^2}(1+\beta)[\operatorname{Tr}(\partial_\mu u^+ \partial^\nu u)]^2. \tag{8}$$

This shows that indeed there are only two independent quartic terms in $\mathcal{L}$.

3. The boundary value problem

By using the hedgehog ansatz

$$u = \exp(i\tau \cdot \hat{r} F(r)), \tag{9}$$

where τ are the Pauli spin matrices and $F(r)$ is the chiral angle. One can obtain from the lagrangian (3), the following differential equation satisfied by $F(r)$:

$$\left[1 + \frac{8\sin^2 F}{\tilde{r}^2}(1-2\gamma) - 12(\beta+2\gamma)F'^2\right]F'' + \frac{2}{\tilde{r}}F' + \frac{4\sin 2F}{\tilde{r}^2}$$
$$\times \left[F'^2(1-2\gamma) - \frac{1}{4} + \frac{\sin^2 F}{\tilde{r}^2}(\beta+4\gamma-1)\right] - \frac{8}{\tilde{r}}(\beta+2\gamma)F'^3 = 0, \tag{10}$$

where $\tilde{r}$ is the dimensions variable $\tilde{r} = eF_\pi r$. Equation (10) has to be solved subject to the boundary conditions

$$F(r) = \pi \quad \text{at } r = 0 \quad \text{and } F(r) = 0 \quad \text{as } r \to \infty. \tag{11}$$

The static mass M of the soliton can be obtained in terms of $F(r)$ and its derivatives:

$$M = 4\pi\left(\frac{F\pi}{e}\right)\int_0^\infty \tilde{r}^2 \left\{\frac{1}{8}\left[F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}\right] + (1+\beta)\frac{\sin^2 F}{2\tilde{r}^2}\left[\frac{\sin^2 F}{\tilde{r}^2} + 2F'^2\right]\right.$$
$$\left. - \frac{1}{4}(\beta+2\gamma)\left[F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}\right]^2\right\} \mathrm{d}\tilde{r}. \tag{12}$$

The solitonic solutions obtained from (9), (10) and (11) are not eigenstates of spin and iso-spin. To obtain solutions with definite spin and iso-spin one makes the transformation $u(t) = A(t)u_0 A^{-1}(t)$ where u_0 is the static solution and $A(t)$ is an arbitrary, time dependent $SU(2)$ matrix. With this transformation, the lagrangian (3) gets transformed to

$$L = \int \mathcal{L} \mathrm{d}^3x = -M + \lambda \operatorname{Tr}[\partial_0 A(t) \partial_0 A^{-1}(t)], \tag{13}$$

where λ is given by

$$\lambda = \frac{4\pi}{6}\frac{1}{e^3 F_\pi}\int_0^\infty \tilde{r}^2 \sin^2 F\left[1 + 4(1+\beta)\left(F'^2 + \frac{\sin^2 F}{\tilde{r}^2}\right)\right.$$
$$\left. - 4(\beta+2\gamma)\left(F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}\right)\right] \mathrm{d}\tilde{r}. \tag{14}$$

The corresponding energy eigenvalues are obtained by projection to the proper spin and iso-spin eigenstates and are given by

$$m_N = M + \frac{3}{8\lambda}, \tag{15a}$$

$$m_\Delta = M + \frac{15}{8\lambda}, \tag{15b}$$

where m_N and m_Δ are nucleon and delta isobar masses respectively.

4. Solutions of the boundary value problem

Before embarking on the numerical solution of the boundary value problem (BVP) defined by equations (10) and (11) it is useful to analyse the nature of the solutions analytically in the limits $r = 0$ and $r \to \infty$. In the vicinity of $r = 0, F(r)$ can be expanded in a power series of the form

$$F(r) = \pi + Cr^n + \cdots. \tag{16}$$

Substituting in eq. (10) one obtains

$$Cr^{n-2}[(n-1)(n+2)] + C^3 r^{3n-4}[-12(\beta+2\gamma)n^4 + 4(\beta+2\gamma)n^3 + 16(1-2\gamma)n^2 - 8(1-2\gamma)n + 8(\beta+4\gamma-1)] + \cdots = 0. \tag{17}$$

For $n = 1$ both the first and second terms in the l.h.s vanish identically. This implies that a power series solution for $F(r)$ in the form $F(r) = \pi + Cr + \cdots$ is possible near $r = 0$ for all values of β and γ.

To study the solutions at $r \to \infty$, eq. (10) is first transformed by the substitution $\xi = 1/r$ and the resulting equation studied at $\xi = 0$. One obtains the following equation:

$$\begin{aligned}&[1 + 8\sin^2 K \cdot \xi^2(1-2\gamma) - 12(\beta+2\gamma)\xi^4 K'^2]\xi^2 K'' \\ &+ 2[8\sin^2 K \cdot \xi^2(1-2\gamma) - 12(\beta+2\gamma)\xi^4 . K'^2]\xi K' \\ &+ 4\sin 2K \cdot [\xi^4 K'^2(1-2\gamma) - \tfrac{1}{4} + \sin^2 K \cdot \xi^2(\beta+4\gamma-1)] \\ &- 8\xi^5(\beta+2\gamma)K'^3 = 0,\end{aligned} \tag{18}$$

where $K(\xi) = F(1/\xi)$.

As before, we try a power series solution of (18) of the form

$$K(\xi) = C'\xi^n + \cdots. \tag{19}$$

Substituting (19) in (18) and equating the coefficient of ξ^n, the lowest power of ξ to zero one obtains

$$C'[n(n-1) - 2] = 0. \tag{20}$$

Equation (20) cannot be satisfied for $n = 0$ and $n = 1$ ($n = 0$ also does not satisfy the boundary condition), but can be satisfied for $n = 2$. This shows that eq. (10) admits of a power series solution of the form $F(r) = C'/r^2 + \cdots$ at $r \to \infty$ irrespective of the values

of β and γ. Thus the asymptotic behaviour of $F(r)$ is identical with those for the Skyrme problem.

The boundary value problem was solved numerically by the shooting procedure for the following cases: (a) $\beta = 0$, $\gamma \neq 0$; (b) $\beta \neq 0$, $\gamma = 0$. It was found that solutions to the BVP exist only for values of the parameters β and γ lying in a range. The stepsize for numerical integration was taken to be 0.1. Since the interval of integration extends from 0 to ∞, a cut off $\tilde{r}_{\max}$ for the upper limit has to be introduced. We have chosen $\tilde{r}_{\max} = 150.1$. This choice of $\tilde{r}_{\max}$ is reasonable as it is about 38 times the pion Compton wavelength and $r^2F(r)$ becomes almost constant in this range of r as is expected from the asymptotic behaviour of $F(r)$ as $r \to \infty$.

(a) *Case* $\beta = 0$, $\gamma \neq 0$. The lagrangian (3) with $\beta = 0$ was investigated by DGH [14] by treating the term proportional to γ as a perturbation. The unperturbed solution of (10) corresponding to $\beta = 0$, $\gamma = 0$, when substituted in (12) and (14) yields [14]

$$M = \left(\frac{73F_\pi}{2e}\right)(1 - 0.77\gamma + \cdots), \tag{21}$$

$$\lambda = \left(\frac{106.6}{e^3F_\pi}\right)(1 - 1.1\gamma + \cdots). \tag{22}$$

By partial wave analysis of the T-matrix for $\pi\pi$ scattering, the parameters F_π, e and γ can be expressed in terms of scattering lengths and slopes [14]. In particular γ is given by

$$\gamma = \frac{1}{4} \cdot \frac{(a_2^0 + 2a_2^2)}{(a_2^0 - a_2^2)}, \tag{23}$$

where a_l^{I} is the scattering length corresponding to partial wave l and iso-spin channel I. Using the experimental values of the $\pi\pi$ scattering parameters, the value of γ is found [14] to be 0.16. M and λ can be obtained from (21) and (22) which, when substituted in (15a) and (15b) yields the values of nucleon and Δ-isobar masses.

Instead of the perturbative approach, we have solved the full boundary value problem (10) and (11) for various values of γ. The chiral angle $F(r)$ obtained this way, when substituted in (12) and (14) gives M and λ for each value of γ. The nucleon and Δ-isobar masses can then be obtained in an analogous manner. It was found that the solution to the boundary value problem (10) and (11) does not exist for $\gamma > 0.11$. In a non-linear boundary value problem, the addition of even a small non-linear term may drastically affect the nature of the solution and perturbative methods may not work.

(b) *Case* $\beta \neq 0, \gamma = 0$. In this case solutions to the BVP exist only for β lying in the range 0 to 0.29. As in the previous case, the parameter β was expressed in terms of $\pi\pi$ scattering parameters by partial wave analysis of the T-matrix. We obtain

$$\beta = \frac{(2a_2^2 + a_2^0)}{(a_2^0 - 4a_2^2)}. \tag{24}$$

Using the experimental values of the scattering lengths quoted in ref. [14], one obtains $\beta = 0.44$. However, the solution to the BVP for this value of β does not exist.

One might suspect that the instabilities encountered in the solution of the BVP for $\gamma > 0.11 (\beta = 0)$ and $\beta > 0.29 (\gamma = 0)$ are due to the behaviour of the static mass under

Derrick's scaling. From eq. (12) the static mass is given by

$$M = M_2 + (1+\beta)M_4 - (\beta + 2\gamma)M_4', \tag{25}$$

where

$$M_2 = 4\pi \cdot \left(\frac{F_\pi}{e}\right) \int_0^\infty \frac{1}{8} \tilde{r}^2 \left[F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}\right] d\tilde{r}, \tag{26}$$

$$M_4 = 4\pi \cdot \left(\frac{F_\pi}{e}\right) \int_0^\infty \frac{1}{2} \sin^2 F \left[\frac{\sin^2 F}{\tilde{r}^2} + 2F'^2\right] d\tilde{r}, \tag{27}$$

and

$$M_4' = 4\pi \cdot \left(\frac{F_\pi}{e}\right) \int_0^\infty \frac{1}{4} \tilde{r}^2 \left[F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}\right] d\tilde{r}. \tag{28}$$

Under scaling $\tilde{r} \to \lambda\tilde{r}$, eq. (25) transforms to

$$M_\lambda = \lambda M_2 + [(1+\beta)M_4 - (\beta + 2\gamma)M_4'] \cdot \frac{1}{\lambda}. \tag{29}$$

When there is no quartic term in the lagrangian (non-linear sigma model), the term in square brackets in (29) is zero and $M = \lambda M_2$. This does not have a minimum with respect to λ, indicating that there is no stable soliton. With Skyrme lagrangian $\beta = \gamma = 0$, $M_\lambda = \lambda M_2 + M_4 \cdot (1/\lambda)$ which has a minimum for $\lambda = \sqrt{M_4/M_2}$. Since M_4 and M_2 are positive definitive, there is always a stable soliton solution for the Skyrme model. With $\beta, \gamma \neq 0$ the term in square brackets in (29) may again vanish for some values of the parameters. When this happens $M_\lambda = \lambda M_2$ and again there is no stable soliton. To check whether the instabilities encountered for $\gamma > 0.11 (\beta = 0)$ and $\beta > 0.29 (\gamma = 0)$ are due to this effect, we have studied the variation of the term in square brackets in (29) as γ approaches 0.11 for $\beta = 0$ and as β approaches 0.29 for $\gamma = 0$. The term accounts for half

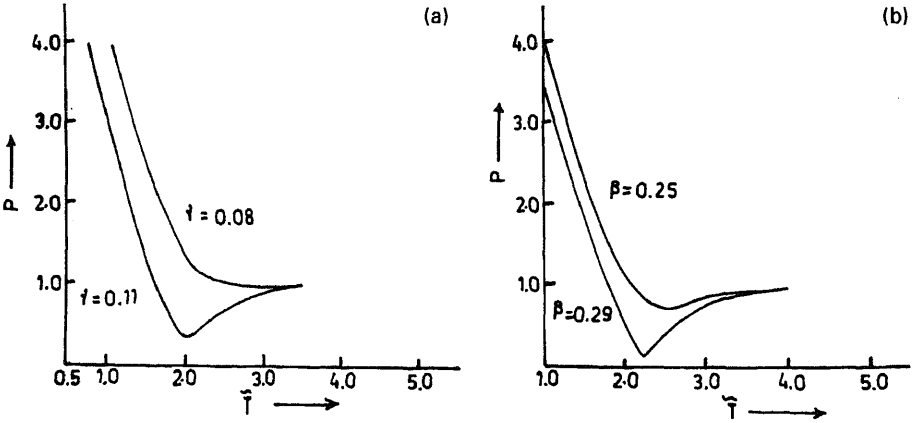


Figure 1. (**a**) and (**b**) show plots of $P = 1 + (8\sin^2 F/\tilde{r}^2)\,(1 - 2\gamma) - 12(\beta + 2\gamma)F'^2$ as a function of $\tilde{r}$ for various values of β and γ, P is the coefficient of F'' in eq. (10).

the static mass and does not approach zero as the upper limits of the parameters are reached indicating that the instabilities are not due to Derrick's scaling behaviour.

The origin of the instabilities can be understood from the differential equation (10). For some values of the parameters β and γ, the coefficient of F'' in (10) may become zero at some point $\tilde{r}$. Then the differential equation, at this particular point becomes first order and the numerical method breaks down. A study of the coefficient of F'' in (10) as a function of β and γ supports this conjecture. Figures 1(a) and 1(b) show plots of the coefficient $P = 1 + (8\sin^2 F/\tilde{r}^2(1-2\gamma) - 12(\beta+2\gamma)F'^2$ of F'' as a function of $\tilde{r}$ for (a) $\gamma = 0.08$ and $\gamma = 0.11$ with $\beta = 0$ and (b) $\beta = 0.25$ and $\beta = 0.29$ with $\gamma = 0$. In both the cases P is positive throughout the interval and has pronounced minima for the limiting values of the parameters β and γ, the minimum values of P being close to zero. Away from the limiting values of β and γ [e.g. for (a) $\gamma = 0.08$, $\beta = 0$ and (b) $\gamma = 0, \beta = 0.25$] the minima are rather flat and the minimum values are higher. It is interesting to note that the coefficient of F'' is never zero for the Skyrme problem.

5. Physical quantities of the nucleon

Various physical quantities of the nucleon can be expressed in terms of the chiral angle $F(r)$ and the parameters F_π, e, β and γ. The analytical expressions are given in ref. [3] for the Skyrme model ($\beta = 0, \gamma = 0$). For $\beta, \gamma \neq 0$, the expressions for iso-scalar mean radius $\langle r^2\rangle^{1/2}_{I=0}$, the iso-scalar magnetic mean radius $\langle r^2\rangle^{1/2}_{M,I=0}$, the proton and neutron magnetic moments are unchanged except that the expressions for M and λ occurring in them are changed as given in (12) and (14). The axial coupling constant g_A is given by

$$g_A = -\frac{\pi}{3e^2} D, \tag{30}$$

where

$$D = \int_0^\infty d\tilde{r}\,\tilde{r}^2 \left[\left(F' + \frac{\sin 2F}{\tilde{r}}\right) + 4\left(\frac{\sin 2F}{\tilde{r}}F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}F' + \frac{\sin^2 F \sin 2F}{\tilde{r}^3}\right)\right.$$
$$-4\beta\left(F'^3 + \frac{\sin^2 F \sin 2F}{\tilde{r}^3}\right) - 8\gamma\left(F'^3 + \frac{\sin 2F}{\tilde{r}}F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}F'\right.$$
$$\left.\left. + \frac{2\sin^2 F}{\tilde{r}^3}\sin 2F\right)\right] \tag{31}$$

and the pion nucleon coupling constant $g_{\pi NN}$ is related to g_A by the relation

$$g_{\pi NN} = \frac{2m_N}{F_\pi} g_A. \tag{32}$$

Table 1 summarizes the calculated values of the physical quantities of the nucleon for $\beta = 0$, $\gamma = 0.11$ and $\beta = 0.29$, $\gamma = 0$ (the highest values for which solution exists). The corresponding values for the Skyrme model and the experimental values are quoted for comparison. It is found that the pion decay constant decreases a little from the Skyrme value making the agreement with experiment slightly worse. But in case of other physical quantities like iso-scalar mean square radius, iso-scalar magnetic mean square radius,

Table 1. Calculated physical quantities of the nucleon for (a) $\beta = 0.29, \gamma = 0$ and (b) $\beta = 0, \gamma = 0.11$. The values of the parameter e in eq. (1) are found to be 4.87 and 4.32 respectively in cases (a) and (b).

Quantity	Our calculation		Skyrme model	Experiment
	$\beta = 0.29, \gamma = 0$	$\beta = 0, \gamma = 0.11$		
Nucleon mass (m_N)	Input	Input	Input	939 Mev
Δ-isobar mass (m_Δ)	Input	Input	Input	1232 Mev
Pion decay constant (F_π)	126.2 Mev	126.6 Mev	129 Mev	186 Mev
Iso-scalar, mean radius $(\langle r^2\rangle^{1/2}_{I=0})$	0.60 fm	0.60 fm	0.59 fm	0.72 fm
Iso-scalar magnetic mean radius $(\langle r^2\rangle^{1/2}_{M,I=0})$	0.86 fm	0.87 fm	0.92 fm	0.81 fm
Proton magnetic moment (μ_p)	1.89	1.89	1.87	2.79
Neutron magnetic moment (μ_n)	−1.32	−1.32	−1.31	−1.91
$\lvert\mu_p/\mu_n\rvert$	1.43	1.43	1.43	1.46
Axial coupling constant (g_A)	0.69	0.67	0.61	1.23
Pion nucleon coupling constant ($g_{\pi NN}$)	10.20	9.96	8.9	13.5

proton magnetic moment, axial coupling constant and pion nucleon coupling constant $g_{\pi NN}$, the agreement with experiment improves as compared to the Skyrme value.

6. Conclusion

We have given a detailed, systematic numerical analysis of the boundary value problem resulting from the most general Skyrme type lagrangian containing quartic terms in field gradients. From our discussion in § 2 it is clear that in the chiral limit, either the lagrangian with $\beta = 0, \gamma \neq 0$ or the one with $\beta \neq 0, \gamma = 0$ can be taken as a suitable quartic generalization of the Skyrme model. In both the cases, solutions to the BVP do not exist for values of the parameters β and γ predicted from pion-pion scattering data. In the case $\beta = 0$, Lacombe *et al* [17] have pointed out that the term proportional to γ in the lagrangian is necessary to produce the right nucleon-nucleon potential in the middle range. With $\beta \neq 0, \gamma = 0$, the lagrangian can be written entirely in terms of commutators and anti-commutators of the current L_μ, a feature which can be useful in the sixth order generalization of the lagrangian.

References

[1] T H R Skyrme, *Proc. R. Soc.* (*London*) **A260**, 127 (1961)
[2] T H R Skyrme, *Nucl. Phys.* **31**, 556 (1962)
[3] G S Adkins, C R Nappi and E Witten, *Nucl. Phys.* **B228**, 552 (1983)
[4] G S Adkins, *Phys. Rev.* **D33**, 193 (1986)
[5] G S Adkins and C R Nappi, *Phys. Lett.* **B137**, 251 (1984)
[6] O Kaymakcalan, S Rajeev and J Schechter, *Phys. Rev.* **D30**, 594 (1984)

[7] U G Meissner and I Zahed, *Phys. Rev. Lett.* **56**, 1035 (1986)
[8] M Mashaal, T N Pham and N Tran Truong, *Phys. Rev. Lett.* **56**, 436 (1986); *Phys. Rev.* **D34**, 3484 (1986)
[9] A Jackson, A D Jackson, A S Goldhaber, G E Brown and L C Castillejo, *Phys. Lett.* **B154**, 101 (1985)
[10] L Marleau, *Phys. Rev.* **D41**, 1606 (1990); **D43**, 885 (1991); **D45**, 1776 (1992)
[11] M Lacombe, B Loiseau and R Vinh Mau, *Phys. Lett.* **B169**, 121 (1986)
[12] B Schwesinger and H Weigel, *Nucl. Phys.* **A465**, 733 (1987)
[13] S Weinberg, *Physica* **96**, 325 (1979)
[14] J F Donoghue, E Golowich and B R Holstein, *Phys. Rev. Lett.* **53**, 747 (1984)
[15] J Gasser and H Leutwyler, *Phys. Lett.* **B125**, 321; 325 (1983)
[16] Hai-Yang Chang, *Phys. Rev.* **D34**, 166 (1986)
[17] M Lacombe, B Loiseau and R Vinh Mau, *Phys. Lett.* **B161**, 31 (1985)

PRAMANA
— journal of
physics

© Printed in India

Vol. 49, No. 4,
October 1997
pp. 431–442

Electromagnetic properties of a chiral-plasma medium

H TORRES-SILVA, C VILLARROEL, P H SAKANAKA* and N REGGIANI*
Departamento de Electrónica, Facultad de Ingeniería, Universidad de Tarapacá, UTA Casilla 6-D, Av. 18 de Septiembre 2222, Arica, Chile
*Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas, 13.083-970 Campina, SP, Brazil

MS received 6 June 1996; revised 16 April 1997

Abstract. The theoretical properties of a composite chiral-plasma medium are developed. By using the reaction theorem for a magnetized chiroplasma, we obtain the proof of nonreciprocity based upon the constitutive relationships between electromagnetic vectors **E**, **B**, **H**, **D**. Using the Maxwell's equations and the proposed constitutive relations for a chiral-plasma medium, we derive the vectors **E** and **H** and from these equations, dispersion relations and **E**-field polarizations are based. The obtained results for waves propagating parallel to the external magnetic field in a cold magnetized chiro-plasma are compared with typical results obtained for a cold plasma. For circulary polarized waves, a new mode conversion is founded with the chiral effect. The chiral rotation is obtained and compared with the Faraday rotation. For waves propagating across the magnetic field, we found a shift of the cut-offs of ordinary and extraordinary waves. On the lower branch of the extraordinary wave mode there is no bands of forbidden frequencies and the reflection point vanishes when the chiral parameter increases.

Keywords. Chiral waves; modes; polarization; chiral-plasma medium; Faraday rotation.

PACS No. 52·35

1. Introduction

Chirality is a geometrical notion which refers to the lack of symmetry of an object. The electromagnetic chirality is also known as optical activity. The handedness of the uniformly distributed and randomly oriented chiral objects, which compose the chiral medium, is responsible for the observed optical activity. Optically active objects are three-dimensional and chiral. A three-dimensional object is chiral if it is not superimposable in its mirror image by translation and rotation.

A chiral medium when interacting with an electromagnetic wave can rotate the plane of polarization of the wave to the right or to the left depending on the handedness of the media. This fact is expected to play an important role in the potential application of chiral media in the microwave and optical regimes.

Chiral media [1, 2] and ferrite media [3] have been studied over the last decade for many applications. Chiral media has been examined as a coating for reducing radar cross section, for antennas and radomes, in waveguides and for microstrip substrates, guided-wave structures and potential application as reflection or antireflection, thin coating and shielding. The principal problem in working with a chiral medium is on the control of the degree of the chirality.

A chiral-plasma medium is examined. The plasma part of the composite medium is non-reciprocal due to an external magnetic field. To find the general dispersion relation, vector phasor Helmholtz based equations are derived, giving the ω against $\mathbf{k}$ behavior. The modal eigenvalue properties in the chiral-plasma medium, which is doubly anisotropic, are determined. We compare our results for the case of waves that propagate parallel to the magnetic field in a cold magnetized chiro-plasma with the typical ones obtained for a cold plasma [4]. For some values of the chirality parameter a new mode conversion appears. Also we obtain the chiral-Faraday rotation which can be compared with the typical Faraday rotation for a pair of right and left-hand circular polarized waves. For waves which propagate perpendicular to the magnetic field, there is no mode conversion but a lower band of forbidden frequencies disappears. Besides we found a shift of the cut-offs of ordinary and extraordinary waves across the external magnetic field.

2. Theoretical foundations

We propose the followings constitutive relations for chiral-plasma media:

$$\begin{aligned} \mathbf{D} &= \vec{\varepsilon}\cdot\mathbf{E} + t_1\mathbf{H}, \\ \mathbf{B} &= \mu\mathbf{H} + t_2\mathbf{E}. \end{aligned} \tag{1}$$

The plasma media constitutive relations are [4]

$$\begin{aligned} \mathbf{D}_p &= \vec{\varepsilon}_p\cdot\mathbf{E}_p, \\ \mathbf{B}_p &= \mu_0\mathbf{H}_p, \end{aligned} \tag{2}$$

where

$$\vec{\varepsilon} = \begin{vmatrix} \varepsilon_1 & i\varepsilon_2 & 0 \\ -i\varepsilon_2 & \varepsilon_1 & 0 \\ 0 & 0 & \varepsilon_3 \end{vmatrix}. \tag{3}$$

Here $\vec{\varepsilon}$, $t_{1,2}$ represents the permittivity tensor and the chirality parameters of the composite medium respectively.

The lossless character of the magnetized cold plasma medium is implied by the Hermitian nature of the tensor $(\vec{\varepsilon}_p^{\,*})^T = \vec{\varepsilon}_p$. The superscripts * and T denote complex conjugate and transpose respectively.

In the search for new media which display nonreciprocal properties, it is essential to establish the nature of the chirality parameter t_1 and t_2.

The anisotropic reaction theorem [5] is

$$\int \mathbf{E}_b\cdot\mathbf{J}_a\,dv = \int \tilde{\mathbf{E}}_a\cdot\mathbf{J}_b\,dv. \tag{4}$$

Here, the sources current $\mathbf{J}_a$, $\mathbf{J}_b$ produce fields $\mathbf{E}_a$, $\mathbf{E}_b$ respectively. The tilde over the fields indicates a modified medium. Thus, we can obtain 6×6 constitutive tensors

$$\tilde{\mathbf{A}} = \begin{vmatrix} \hat{\varepsilon}^T & -\hat{t}'^T \\ -\hat{t}^T & \hat{\mu}^T \end{vmatrix} \tag{5}$$

and

$$\hat{\mathbf{A}} = \begin{vmatrix} \hat{\varepsilon} & -\hat{t} \\ -\hat{t} & \hat{\mu} \end{vmatrix} \tag{6}$$

with $\hat{t}$ and $\hat{t}'$ being the optical activity 3×3 tensors.

Reciprocity occurs only if

$$\int \mathbf{E}_b \cdot \mathbf{J}_a \, dv = \int \mathbf{E}_a \cdot \mathbf{J}_b \, dv. \tag{7}$$

By (4) it requires

$$\tilde{\mathbf{A}} = \hat{\mathbf{A}}. \tag{8}$$

For chiral media we must obtain

$$\hat{\varepsilon} = \vec{\varepsilon}, \ \hat{t} = t_1 I, \ \hat{t}' = t_2 I, \ \hat{\mu} = \mu, \tag{9}$$

where I is the unity tensor. To obtain reciprocity (8) imposes

$$-t_2 I^{\mathrm{T}} = t_1 I, \ -t_1 I^{\mathrm{T}} = t_2 I. \tag{10}$$

That is

$$t_1 = -t_2. \tag{11}$$

In the case of plasma media equations (2) hold, leading to

$$\hat{\varepsilon} = \vec{\varepsilon}_{\mathrm{p}}, \ \hat{t} = \hat{t}' = 0, \ \hat{\mu} = \mu_0, \tag{12}$$

Then, for the proposed constitutive relations, eq. (1), we have

$$\mathbf{D} = \vec{\varepsilon}_{\mathrm{p}} \cdot \mathbf{E} + t_1 \mathbf{H}, \tag{13}$$

$$\mathbf{B} = \mu_0 \mathbf{H} + t_2 \mathbf{E}. \tag{14}$$

A complete study that characterize electromagnetic waves in general bianisotropic media was developed by Kong [6] and our approach is a particular case.

3. Vector Helmholtz equations

The E-field vector Helmholtz equations is derived by inserting the constitutive relation, (13) and (14), into Maxwell's equations

$$\nabla \times \mathbf{E} = -i\omega \mathbf{B}, \tag{15}$$

$$\nabla \times \mathbf{H} = i\omega \mathbf{D} + \mathbf{J}, \tag{16}$$

so

$$\nabla \times \mathbf{E} = -i\omega\mu_0 \mathbf{H} - i\omega t_2 \mathbf{E}, \tag{17}$$

$$\nabla \times \mathbf{H} = i\omega \vec{\varepsilon} \cdot \mathbf{E} + i\omega t_1 \mathbf{H}. \tag{18}$$

Solving for $\mathbf{H}$, (17) gives

$$\mathbf{H} = \frac{1}{\mu_0}\left(\frac{i}{\omega}\nabla \times \mathbf{E} - t_2 \mathbf{E}\right) \tag{19}$$

and putting this into (18) we obtain

$$\nabla \times \mathbf{H} = \frac{1}{\mu_0}\frac{i}{\omega}(\nabla \times \nabla \times \mathbf{E}) - \frac{t_2}{\mu_0}\nabla \times \mathbf{E}. \tag{20}$$

Then the **E** field vector equation becomes

$$\nabla \times \nabla \times \mathbf{E} + i\omega(t_2 - t_1)\nabla \times \mathbf{E} - \omega^2 \mu_0 \varepsilon_0 \left(\frac{\overleftrightarrow{\varepsilon}}{\varepsilon_0} - \frac{t_1 t_2}{\mu_0 \varepsilon_0} \right) \cdot \mathbf{E} = 0. \tag{21}$$

Here the plasma current is included in the permittivity tensor $\overleftrightarrow{\varepsilon}$. Similarly the **H** field vector equation is

$$\nabla \times \overleftrightarrow{\varepsilon}^{-1} \nabla \times \mathbf{H} + i\omega(t_2 \overleftrightarrow{\varepsilon}^{-1} \nabla \times \mathbf{H} - t_1 \nabla \times \overleftrightarrow{\varepsilon}^{-1} \times \mathbf{H}) - \omega^2 \mu_0 \left(I - \frac{t_1 t_2}{\mu_0} \overleftrightarrow{\varepsilon}^{-1} \right) \mathbf{H} = 0. \tag{22}$$

The inverse permittivity tensor is

$$\overleftrightarrow{\varepsilon}^{-1} = \begin{vmatrix} \varepsilon_1 & -i\varepsilon_2 & 0 \\ i\varepsilon_2 & \varepsilon_1 & 0 \\ 0 & 0 & \dfrac{\varepsilon_1^2 - \varepsilon_2^2}{\varepsilon_3} \end{vmatrix} \Bigg/ (\varepsilon_1^2 - \varepsilon_2^2).$$

4. Dispersion relation

The dispersion relation for the propagation vector **k** against ω can be obtained from **E** or **H** vector equations. We start with the **E** field relation which is simpler than **H** vector equation.

Setting **E** as follows,

$$\mathbf{E} = \mathbf{E}_0 e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{23}$$

we obtain

$$-\mathbf{k} \times \mathbf{k} \times \mathbf{E}_0 + \omega(t_2 - t_1)\mathbf{k} \times \mathbf{E}_0 - \omega^2 \mu_0 \varepsilon_0 \left(\frac{\overleftrightarrow{\varepsilon}}{\varepsilon_0} - \frac{t_1 t_2}{\mu_0 \varepsilon_0} \right) \mathbf{E}_0 = 0. \tag{24}$$

Putting $\mathbf{E}_0$ into rectangular coordinates

$$\mathbf{E}_0 = E_x \hat{x} + E_y \hat{y} + E_z \hat{z}, \tag{25}$$

we obtain three component systems of equations which determine the eigenvector. The determinant of the coefficient component matrix M_k will determine the eigenvalues, thereby yielding the ω against **k** dispersion diagram in phase-space.

Writing det $M_k = 0$, with $k_x = 0$ and with symmetry about the z-axis we obtain

$$\begin{vmatrix} 1 - \dfrac{\varepsilon_1}{n^2 \varepsilon_0}\left(1 - \dfrac{t_1 t_2}{\mu_0 \varepsilon_1}\right) & -\dfrac{i\varepsilon_2}{n^2 \varepsilon_0} - \dfrac{\cos\theta(t_2 - t_1)}{n\sqrt{\mu_0 \varepsilon_0}} & \dfrac{\sin\theta(t_2 - t_1)}{n\sqrt{\mu_0 \varepsilon_0}} \\ \dfrac{i\varepsilon_2}{n^2 \varepsilon_0} + \dfrac{\cos\theta(t_2 - t_1)}{n\sqrt{\mu_0 \varepsilon_0}} & \cos^2\theta - \dfrac{\varepsilon_1}{n^2 \varepsilon_0}\left(1 - \dfrac{t_1 t_2}{\mu_0 \varepsilon_1}\right) & -\sin\theta\cos\theta \\ -\dfrac{\sin\theta(t_2 - t_1)}{n\sqrt{\mu_0 \varepsilon_0}} & -\sin\theta\cos\theta & \sin^2\theta - \dfrac{\varepsilon_3}{n^2 \varepsilon_0}\left(1 - \dfrac{t_1 t_2}{\varepsilon_3 \mu_0}\right) \end{vmatrix} = 0. \tag{26}$$

Here the refractive index n is defined as $n = ck/\omega$, $c = 1/\sqrt{\mu_0 \varepsilon_0}$.

If $\mu_0 = 1, \varepsilon_0 = 1, t_1 = t_2 = 0$, we obtain the same result given by Krall and Trivelpiece for a magnetized plasma [4].

For a collisionless chiroplasma, using (11) we have $t_1 = -it\sqrt{\mu_0\varepsilon_0}$, $t_2 = it\sqrt{\mu_0\varepsilon_0}$, the nontrivial solution of this system comes from setting the determinant of the coefficients equal to zero giving

$$\omega = \omega(n^2, \omega_p, \omega_c, \varepsilon_1, \varepsilon_2, t\ \varepsilon_3, \theta, k). \tag{27}$$

Equation (27) is then the general dispersion relation for waves propagating in a cold collisionless homogeneous chiroplasma in a uniform magnetic field. For a given plasma frequency ω_p, cyclotron frequency ω_c, wave frequency ω and the direction of propagation θ, equation (24) can be solved for the index of refraction n having as parameter the chirality t. From here, for the permittivities we will use relative values. In terms of k, the dispersion relation is

$$a_1k^4 + a_2k^3 + a_3k^2 + a_4k + a_5 = 0, \tag{28}$$

where

$$a_1 = -\left[\omega^2\varepsilon_1\left(1 - \frac{t^2}{\varepsilon_1}\right) - \omega^2\varepsilon_3\left(1 - \frac{t^2}{\varepsilon_3}\right)(\sin^2\theta - 1)\right],$$

$$a_2 = 0,$$

$$a_3 = \omega^2\varepsilon_1\left(1 - \frac{t^2}{\varepsilon_1}\right)\omega^2\varepsilon_3\left(1 - \frac{t^2}{\varepsilon_3}\right)(2 - \sin^2\theta) + \left[\omega^2\varepsilon_1\left(1 - \frac{t^2}{\varepsilon_1}\right)\right]^2$$

$$-(\omega^2\varepsilon_2)^2\sin^2\theta - 4\omega^2t^2\left[\omega^2\varepsilon_1\left(1 - \frac{t^2}{\varepsilon_1^2}\right)\right]\sin^2\theta + \omega^2\varepsilon_3\left(1 - \frac{t^2}{\varepsilon_3}\right)\cos^2\theta,$$

$$a_4 = -2i\omega^2\varepsilon_2\omega^2\varepsilon_3\left(1 - \frac{t^2}{\varepsilon_3}\right)(-2it\omega)\cos\theta,$$

$$a_5 = \omega^2\varepsilon_1\left(1 - \frac{t^2}{\varepsilon_1}\right)\left[\omega^4\varepsilon_2^2 - \omega^4\varepsilon_1^2\left(1 - \frac{t^2}{\varepsilon_1}\right)^2\right].$$

Here, there are four different eigenmodes for **k** implied by (28)

$$\varepsilon_1 = 1 - \frac{\omega_p^2}{\omega^2 - \omega_c^2}, \tag{29}$$

$$\varepsilon_2 = -\frac{\omega_c}{\omega}\frac{\omega_p^2}{\omega^2 - \omega_c^2}, \tag{30}$$

$$\varepsilon_3 = 1 - \frac{\omega_p^2}{\omega^2}, \tag{31}$$

where ω_p is the plasma frequency and ω_c is the gyrofrequency given by

$$\omega_p^2 = \frac{4\pi n_e e^2}{m_e}, \tag{32}$$

$$\omega_c = \frac{eB_0}{m_e c}. \tag{33}$$

Setting $\theta = 0$, it is posible from eq. (21) different wave modes, writing the E-field vector equation in the form

$$(n^2 - \varepsilon_R) E_R = 0,$$

$$(n^2 - \varepsilon_L) E_L = 0,$$

$$\varepsilon_3 \left(1 - \frac{t^2}{\varepsilon_3}\right) E_z = 0. \tag{34}$$

From these equations, the last one gives a longitudinal mode. When $\varepsilon_3 = t^2$, we obtain the longitudinal electron plasma oscillations modified by the chiral parameter t. Since there is no wave propagation along the magnetic field, these chiroplasma oscillations do not constitute a propagation mode. For the other modes we have

$$\varepsilon_{R,L} = \varepsilon_1 \left(1 - \frac{t^2}{\varepsilon_1}\right) \pm \varepsilon_2 \left(1 + \frac{2tn}{\varepsilon_2}\right) \tag{35}$$

and

$$E_{R,L} = E_x \pm iE_y. \tag{36}$$

It is useful to explore these solutions in terms of the wave numbers k

$$k_R = \frac{t\omega}{c} \pm \frac{\omega}{c}\sqrt{\varepsilon_1 + \varepsilon_2} \tag{37}$$

$$k_L = -\frac{t\omega}{c} \pm \frac{\omega}{c}\sqrt{\varepsilon_1 - \varepsilon_2}, \tag{38}$$

where k_R is the wave number for a circularly polarized wave which drive electrons in the direction of their cyclotron motion, i.e., right-hand circularly polarized waves (RCP) and k_L is the wave number for a circularly polarized wave which drive electrons in the direction opposite of their cyclotron motion, i.e., left-hand circularly polarized waves (LCP). The t parameter modifies the typical plot of $\omega(k)$ shown by Krall and Trivelpiece, where the cutoff frequencies are shifted. Also the reflection points of the RCP and LCP are shifted. However the resonance which occurs when the wave phase velocity goes to zero is not modified by the chiral parameter. In figure 1 we present the modifications introduced by the parameter t in the dispersion relations of the right and left polarized waves. In this figure, the dispersion relations of the right and left circularly polarized waves are indicated by circles and stars, respectively. When $t \neq 0$, ε_1 and ε_3 depend on t and k_R and k_L have a linear term, $t\omega/c$, as can be seen in (37) and (38). In this way instead of modifying the curves that exist for $t = 0$, the parameter t permits the wave to propagate in a region of frequencies that is forbidden in the case $t = 0$. The RCP wave mode, in the lower branch, also known as the electron cyclotron wave is weakly modified by the chiral parameter, but the upper branch is strongly shifted when t increases. We can observe in figure 1 that for $t = 0$, there is no intersection of the dispersion relations of the right and left circularly polarized waves. Here the frequency bands for which there is no wave propagation can be identified in the plots. When $t \neq 0$ we can observe that there is an intersection of these curves, indicating that the presence

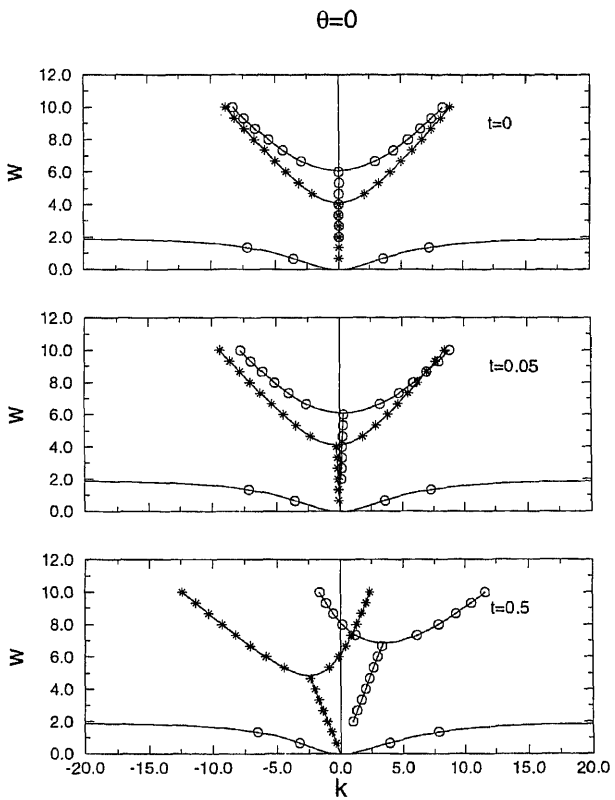


Figure 1. Dispersion relations for various values of the parameter t when the direction of propagation is parallel to the magnetic field ($\theta = 0$). The curves indicated by circles and stars correspond to the right and left circularly polarized waves, respectively.

of the t parameter permits that a wave changes its polarization. In figure 1, for $t = 0$ we can also observe that there is a region where only right circularly polarized waves propogate, a region where only left circularly polarized waves propagate and a region where both propagate. If their amplitudes are equal, the effect of the superposition of a left and right circularly polarized wave is to produce a plane wave with a particular plane of polarization. Because the two polarizations propagate at different velocities, the plane of polarization rotates as the wave propagates along the magnetic field. This effect is called Faraday-chiral rotation [7].

The global rotation of a plane of polarization as a function of distance in the direction of propagation is given by

$$\frac{E_x}{E_y} = \cot\left(\frac{k_L - k_R}{2}\right) z, \tag{39}$$

which means, the presence of the t parameter also affects the Faraday rotation. This chiral-Faraday rotation can be used as a plasma probe. In laboratory experiment this would be done by launching a planewave along the magnetic field in a chiroplasma.

Considering that the plane of polarization of this wave can be determined by an antenna and that we know the magnetic field, the density of the plasma and the frequency of the launched wave, the measurement of the plane of polarization away from the source can determine the value of the parameter t. For instance, considering the plasma frequency, $\omega_p = 5 \cdot 10^7\,s^{-1}$, for the electron gyrofrequency, $\omega_c = 2 \cdot 10^7\,s^{-1}$, and for the launched wave, $\omega = 6 \cdot 5 \cdot 10^7\,s^{-1}$, the value of E_x/E_y, 1 cm away from the source is $E_x/E_y = 15 \cdot 38$, $E_x/E_y = 11 \cdot 48$ and $E_x/E_y = 9 \cdot 15$ for $t = 0$, $t = 0 \cdot 05$ and $t = 0 \cdot 1$, respectively. The chiral rotations are 3·72, 4·97 and 6·24 degrees respectively. Another important effect caused by the presence of the parameter t is the conversion of modes [8]. In figure 1 for $t = 0 \cdot 5$, we can observe that there is a region where both RCP and LCP propagate, for $k = 1 \cdot 0$ we have a mode conversion from RCP to LCP wave. The explicit expression for the frequency around which intersection takes place can be obtained from $(\varepsilon_2^*)^2 = 4t^2(\varepsilon_1)^*$, where the quantitites with $*$ are modified forms of those of a collisionless biased plasma [9], that is, $\varepsilon_1^* = \varepsilon_1 - t^2$, $\varepsilon_3^* = \varepsilon_3 - t^2$ and $\varepsilon_2^* = \varepsilon_2$. This means that the energy of the electrons, obtained from the RCP electromagnetic wave at the electron cyclotron frequency can be transferred to the ions and this mode conversion can be used as a means of heating of the plasma. Note that the RCP wave rotates in the same direction as the electrons about the magnetic field and near the resonance, the energy is transferred from the wave field to the electrons but the mode conversion allows the absorption of energy by the ions.

In connection with the mode conversion and crossover frequencies in plasma, we note that a plasma with only two species may also show lack of symmetry, change of polarization and intermode coupling in the presence of the coriolis force. A detailed study of wave propagation in a rotating cold plasma including the coriolis force has been made by Uberoi and Das [10, 11]. Our results correspond to the very high frequency limit where the rotation and dynamic of ions are not taken into account. As the chiral effect is important at microwave and optical frequencies, if we extend the study by considering the Uberoi and Das [10, 11] approach, we can find that the result of these authors on the polarization reversal, intermode coupling of magnetic and inertial mode in a rotating magnetoplasma are not modified essentially.

In terms of the index of refraction, all the waves supported by the medium can be obtained from the following matrix if n_x and n_y are set equal to zero:

$$\mathbf{M} = \begin{bmatrix} n^2 - \varepsilon_1^* & -2itn - i\varepsilon_2^* & 0 \\ 2itn + i\varepsilon_2^* & n^2 - \varepsilon_1^* & 0 \\ 0 & 0 & \varepsilon_3^* \end{bmatrix} \tag{40}$$

where $t_1 = -it\sqrt{\mu_0 \varepsilon_0}$, $t_2 = it\sqrt{\mu_0 \varepsilon_0}$, $\varepsilon_3^* = \varepsilon_3 - t^2$, $\varepsilon_1^* = \varepsilon_1 - t^2$, $\varepsilon_2^* = \varepsilon_2$, $\mu_0 = 1$, $\varepsilon_0 = 1$. The quantities with $*$ are modified forms of those of a collisionless biased plasma [9].

Solving $\det M_n = 0$, under these conditions, it is possible to obtain

$$\begin{aligned} n_{\parallel p+} &= t + \sqrt{\varepsilon_1 + \varepsilon_2} = t + \sqrt{\mathrm{R}}, \\ n_{\parallel p-} &= t - \sqrt{\varepsilon_1 + \varepsilon_2} = t - \sqrt{\mathrm{R}}, \\ n_{\parallel a+} &= -t + \sqrt{\varepsilon_1 - \varepsilon_2} = -t + \sqrt{\mathrm{L}}, \\ n_{\parallel a-} &= -t - \sqrt{\varepsilon_1 - \varepsilon_2} = -t - \sqrt{\mathrm{L}}. \end{aligned} \tag{41}$$

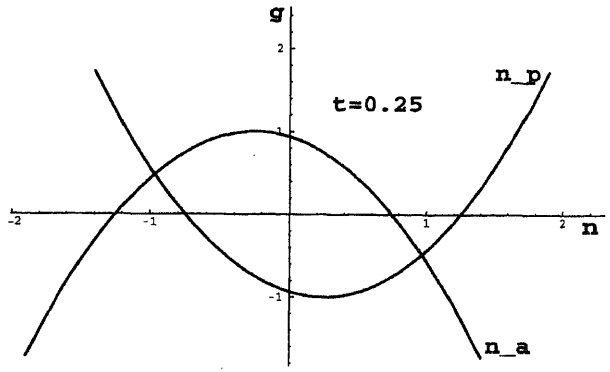


Figure 2. A graph of the $g = \varepsilon_2$ versus n representing the four wavenumbers along the magnetostatic field of a chiroplasma, the chiral parameter $t = 0{\cdot}25$.

Therefore, four wavenumbers are found which are dependent on the permittivity and chirality parameters of the composite medium.

The subscripts 'p' and 'a' refer to the parallel and antiparallel directions of energy propagation, that is the direction of the real part of the Poynting's vector with respect to the static magnetic field, while the plus (R) and minus (L) subscripts denote right-circular polarized (RCP) and left-circular polarized (LCP) propagating waves. For $t = 0{\cdot}25$, figure 2 show the $g = \varepsilon_2$ dependence of the four solutions of the index of refraction present along the magnetostatic field of a chiroplasma. It is possible to compare these results with those obtained by Engheta *et al* [7], in which the description of the Faraday chiral media is obtained using the chirality admittance. Here similar results are obtained in a simpler manner [12].

The helicity and polarization state corresponding to each of the wave numbers can be found by substituting (39) into the condition that the electric field must satisfy:

$$\overleftrightarrow{M} \cdot \mathbf{E} = 0. \tag{42}$$

We note that $n_{\|p+}$ and $n_{\|a+}$ are of positive helicity while $n_{\|p-}$ and $n_{\|a-}$ are of negative helicity.

Also the wave impedances of the positive and negative helicities, η_1 and η_2, respectively, are found to be

$$\eta_1 = \frac{1}{\sqrt{(\varepsilon_1^* + \varepsilon_2)}}, \tag{43}$$

$$\eta_2 = \frac{1}{\sqrt{(\varepsilon_1^* - \varepsilon_2)}}. \tag{44}$$

6. Wave propagation across the magnetic field

By making $\theta = 90°$ the solution to $\det M = 0$ yields a wave with elliptical polarization in the plane transverse to the direction of propagation. The angle θ is defined as the

dispersion relation.

$$k_{\perp\text{extr.}} = \pm \frac{\sqrt{A - \sqrt{B}}}{\sqrt{2(\varepsilon_1 - t^2)}} \tag{45}$$

and

$$k_{\perp\text{ord.}} = \pm \frac{\sqrt{A + \sqrt{B}}}{\sqrt{2(\varepsilon_1 - t^2)}}, \tag{46}$$

where

$$A = \frac{\omega^2}{c^2}[\varepsilon_1^2 - \varepsilon_2^2 + \varepsilon_1\varepsilon_3 + t^2(\varepsilon_1 - \varepsilon_3) - 2t^4] \tag{47}$$

and

$$\begin{aligned} B = \frac{\omega^4}{c^4}\big[&\big((\varepsilon_1^2 - \varepsilon_2^2) - \varepsilon_1\varepsilon_3\big)^2 \\ &+ t^2(6\varepsilon_1^3 - 6\varepsilon_1\varepsilon_2^2 - 2\varepsilon_1\varepsilon_3^2 + 12\varepsilon_1^2\varepsilon_3 - 2\varepsilon_2^2\varepsilon_3) \\ &+ t^4(-15\varepsilon_1^2 + 8\varepsilon_2^2 - 18\varepsilon_1\varepsilon_3 + \varepsilon_3^2) \\ &+ 8t^6(\varepsilon_1 + \varepsilon_3)\big]. \end{aligned} \tag{48}$$

It should be pointed out that the electric field of the extraordinary wave, $k_{\perp\text{extr.}}$ is perpendicular to the magnetic field and the electric field of the ordinary wave, $k_{\perp\text{ord.}}$ is parallel to the magnetic field. In figure 3 we present the effect of the parameter t on the dispersion relations for the case $\theta = \pi/2$. In this figure the ordinary and extraordinary waves are indicated by circles and stars, respectively. When $t = 0{\cdot}05$, for $\theta = \pi/2$, the effect of the parameter is very small. We can observe that the dispersion relations are a little modified, but the parameter is not able to break up the forbidden regions that exist when $t = 0$. When $t = 0{\cdot}5$, the dispersion relations present very different curves with respect to the curves for $t = 0$. The difference in the way the t parameter acts in the parallel and perpendicular directions is due to the kind of equations we have. In (37) and (38) the t parameter appears as a linear term and in (45) and (46) the t parameter appears just inside a square root [6]. Also, we see that for $\theta = \pi/2$ the parameter t does not lead to the conversion of modes, as happens when $\theta = 0$. However, at $t = 0{\cdot}5$, the dispersion relation for the ordinary wave mode is flat and the reflection point is shifted. Now the ordinary and extraordinary wave mode do not have the same value of ω when k is very large. The more important effect is on the lower branch of the extraordinary wave mode because there is no more bands of forbidden frequencies, and the reflection point vanishes.

To see the effects of the chirality parameter on the cut-off frequencies, we can reduce the factors A and B in terms of the Stix's parameters P, R, and L [13], as

$$n_{\perp\text{ord.}} = \pm \sqrt{P\left(1 + \frac{2t^2}{R+L}\right)}, \tag{49}$$

$$n_{\perp\text{extr.}} = \pm \sqrt{\frac{2RL}{R+L}\left(1 - \frac{2t^2}{R+L}\right)}. \tag{50}$$

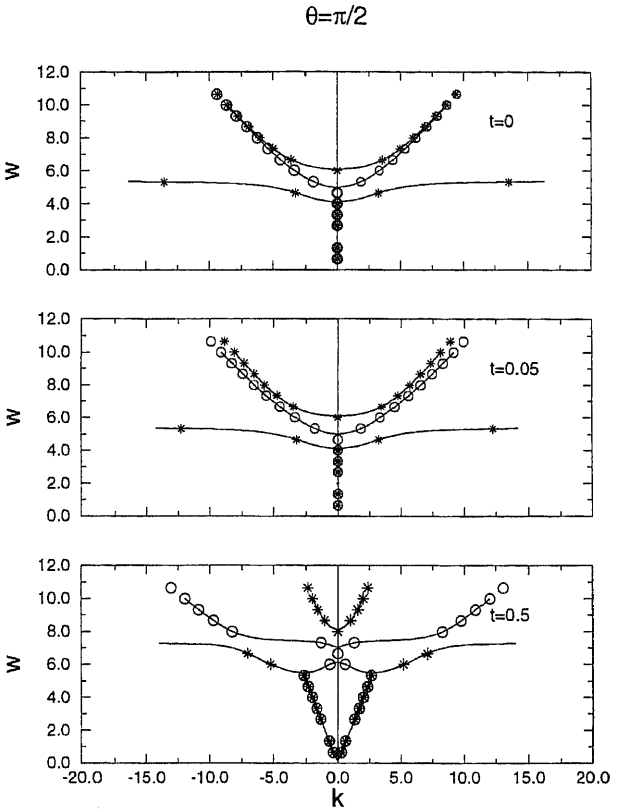


Figure 3. Dispersion relations for various values of the parameter t when the direction of propagation is perpendicular to the magnetic field ($\theta = \pi/2$). The curves indicated by circles and stars correspond to the ordinary and extraordinary waves, respectively.

As t goes to zero we obtain the typical results for a collisionless magnetized plasma. It follows from the equations that the cold plasma cut-offs $P = 0$, $R = 0$ and $L = 0$ are shifted for a chiroplasma in the domain $(\omega_c^2/\omega^2, \omega_p/\omega)$. The respective localization of the lines are schematically presented in figure 4, the vertical dashed lines correspond to $P = 0$.

7. Conclusion

We have examined the problem of high frequency waves in a magnetized chiral plasma. In the case of waves which propagate parallel to the magnetic field, we found a new mode conversion as the chiral parameter increases. This phenomenon is similar to polarization reversal and intermode coupling of slow and fast mode of propagation at low-frequency wave in a rotating magnetoplasma studied by Uberoi and Das. Here we obtain a combined effect Faraday-chiral rotation. Also, we found a shift in the cut-offs of ordinary and extraordinary waves across the external magnetic field.

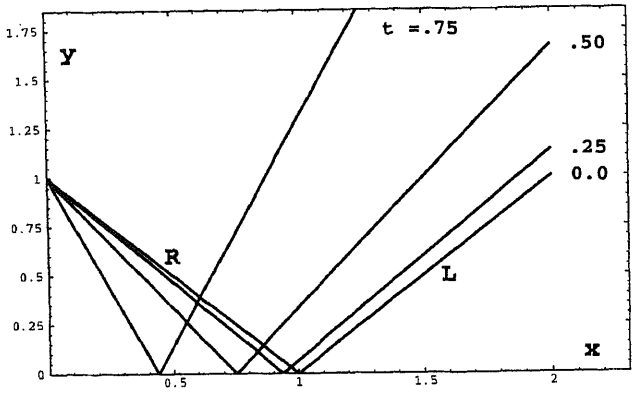


Figure 4. The lines $P = 0$, $R = 0$ and $L = 0$ in the plane $(\omega_c^2/\omega^2, \omega_p/\omega)$ corresponding to the cut-offs in a cold plasma ($t = 0$) and chiroplasma for some values of t.

Acknowledgements

The authors are grateful for financial supports from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil), Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP, Brazil) and Comision Nacional de Ciencia y Tecnologia (Fondecyt, grant 1950963, Chile).

References

[1] D L Jaggard, J C Liu and X Sun, *Electron Lett*, **77**, 27 (1991)
[2] C M Krowne, *IEE Proc.* **140**, 242 (1993)
[3] N Engheta and P Pelet, *Opt. Lett.* **593**, 14 (1989)
[4] N A Krall and A K Trivelpiece, *Principles of plasma physics* (McGraw-Hill Inc., 1973)
[5] C M Krowne, *IEEE Trans. Antennas and Propag.* **24**, 12 (1984)
[6] J A Kong, *Proc. IEEE*, **60**, 1036 (1972)
[7] N Engheta, D L Jaggard and M K Kovarz, *IEEE Trans. Antennas and Propagation* **40**, 367 (1992)
[8] H Torres Silva, P H Sakanaka and N Reggiani, *Waves propagation in a chiral-plasma-media* (CB mag' 95–Congresso Brasileiro de Electromagnetismo, Brasil, 1995) vol. I, p. 201
[9] C H Papas, *Theory of electromagnetic wave propagation* (New York, McGraw-Hill, 1965)
[10] C Uberoi and G C Das, *Plasma Phys.* **14**, 201 (1972)
[11] C Uberoi, *Phys. Fluids* **19**, 921 (1976)
[12] H Torres Silva, P H Sakanaka and N Reggiani, *Rev. Mex. de Fisica* **42**, 101 (1996)
[13] T H Stix, *The theory of plasma waves* (New York: McGraw-Hill, 1962)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 4,
October 1997
pp. 443–453

Bulk viscosity of neutron stars

V K GUPTA, A WADHWA*, S SINGH and J D ANAND
Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India
*Department of Physics, S.G.T.B. Khalsa College, Delhi 110007, India

MS received 17 October 1996; revised 5 February 1997

Abstract. Viscosity of neutron stars has been a continuing area of research for many years now. Recently interest in this field has revived because of the possibility of URCA processes in neutron stars. In this paper we report calculation of the bulk viscosity of neutron stars from these processes. For this purpose we have used the β-decay rates which were calculated without making the usual approximations of neglecting the neutrino momentum and using the nuclear mean field theory for the description of interacting nuclear matter. Also we have not restricted our calculation to the linear regime which corresponds to the assumption that fluctuations in the chemical potential away from β-equilibrium remain small: $\Delta\mu/kT \ll 1$. We find that for large amplitude fluctuations, where the linear approximation is not valid, bulk viscosity increases by many orders of magnitude. Also, as against strange matter stars, where the viscosity first increases with increasing temperature and then starts decreasing beyond 0.1 MeV, we find that the viscosity increases uniformly with temperature at least up to 2 MeV. We discuss the implications of these results for the stability of neutron stars.

Keywords. Bulk viscosity; radial damping; stability of neutron stars.

PACS Nos 21.65; 95.30; 97.60

1. Introduction

The viscosity of neutron star matter has been studied for more than a decade now and it still remains an active area of research [1]. Viscosity determines the damping time scales of radial vibrations of neutron stars [2] which may have been excited at the time of their formation from supernovae explosions. Viscosity also determines the criteria for the gravitational wave instabilities in rapidly rotating neutron stars and thereby the maximum rates of their rotation [3].

Another area of ongoing research is that of strange quark matter. Matter composed of comparable number of u, d and s quarks is conjectured to be the stable (or meta–stable) state of matter [4, 5]; if so, some of the neutron stars may turn out to be strange stars (composed of strange matter) or hybrid stars i.e., those having strange matter cores enveloped by ordinary nuclear matter.

Observationally, it is not easy to distinguish strange matter, hybrid and ordinary neutron stars since in the observed mass region of around 1.4 times the solar mass, all the three have similar radii. The neutrino cooling rates, considered to be one of the best candidates for this purpose, have lost much of their discriminating power since the revival

of ordinary URCA processes in nuclear matter has narrowed down the difference in the expected neutrino cooling rates from the two types of matter [6, 7].

Many studies of the viscosities of the two types of matter have been reported in literature [8–13]. Recently there has been a revival of interest in this problem mainly because of the possibility of the direct URCA processes in nuclear matter. Consequently, the values of quantities determining nuclear transport properties are bound to change considerably.

The bulk viscosity of neutron star matter from direct URCA processes viz,

$$\begin{aligned} n &\rightarrow p + e + \bar{\nu}_e \\ p + e &\rightarrow n + \nu_e \end{aligned} \tag{1}$$

have been calculated by Haensel and Schaeffer [10]. The source of bulk viscosity of neutron star matter is the deviation from β-equilibrium and the ensuing non-equilibrium reactions implied by the compressions and rarefactions of the matter in the pulsating neutron star; thus they are driven by the non-zero value of $\Delta = \mu_n - \mu_p - \mu_e$, away from β-equilibrium. Haensel and Schaeffer [10] assumed that $\Delta/kT \ll 1$. As has been pointed out by Madesen [12], this assumption is certainly not correct at low temperatures. It is also not expected to be valid far away from equilibrium. In the context of the viscosity of strange quark matter Madesen [12] and Goyal *et al* [13] have shown that relaxation of this condition could lead to very significant modifications in the results. Moreover, Haensel and Schaeffer have made the usual simplifying assumptions [14] regarding the rates of various β-decay processes involved. In particular, the angular integrals involved in the calculation of the matrix element are evaluated only approximately and the neutrino momentum in the momentum-conserving δ-function is ignored. These assumptions though necessary to obtain the results in a simple analytic form can, however, alter the rates significantly [7, 15]. This has been demonstrated both for strange [15] as well as nuclear matter [7]. In view of the importance of bulk viscosity in the damping of radial vibrations, in determining the maximum rotation rates of the neutron stars and as means of observationally distinguishing between strange and neutron stars, we feel that a more accurate calculation of this quantity without these simplifying assumptions is in order.

In this paper we report the calculation of the bulk viscosity of neutron stars from the URCA processes (1). For this purpose we have not made use of the linear approximation viz., $\Delta\mu \ll kT$ which is valid for small perturbations only. The direct URCA processes in dense nuclear matter are possible only if the proton fraction $x = n_p/n_B$ is greater than a critical value of $1/9$. Lattimer *et al* [6] have shown that there are simple models of nuclear matter where x is indeed greater than this critical value. Following the revival of interest in the URCA processes [6], we calculated the improved rates of energy loss by neutrino emission from neutron stars [7], without making the usual approximation of neglecting the neutrino momentum. We have followed the same approach for calculation of the rates of various β-decay processes involved. For the description of dense nuclear matter, we have used Walecka's mean field theory [16–18] (including the ρ-meson contribution), as extended to include strong interactions among the particle species involved. As is well-known this theory gives a fairly good description of strong interactions among the various particles and can be easily extended to include higher resonances as well. In § 2 we derive the expressions for bulk viscosity; in § 3 we present our results followed by a brief discussion.

In this section we derive an expression for the bulk viscosity of a neutron star using an approach similar to that of Wang and Lu [8], Sawyer [11] and Madesen [12]. Let there be a periodic fluctuation in the specific volume v of the star according to the following relation

$$v(t) = v_0 + \Delta v \sin(2\pi t/\tau) = v_0 + \delta v(t), \tag{2}$$

where v_0 is the equilibrium volume and Δv the amplitude of the perturbation. The mean dissipation rate of energy per unit mass can be expressed as

$$(\mathrm{d}w/\mathrm{d}t)_{\mathrm{av}} = (-1/\tau)\int_0^\tau P(t)(\mathrm{d}v/\mathrm{d}t)\mathrm{d}t \tag{3}$$

where the pressure $P(t)$ can be expressed near its equilibrium value P_0 as

$$P(t) = P_0 + (\partial P/\partial v)_0 \delta v + (\partial P/\partial n_p)_0 \delta n_p + (\partial P/\partial n_e)_0 \delta n_e + (\partial P/\partial n_n)_0 \delta n_n. \tag{4}$$

The change in the number of neutrons, protons and electrons per unit mass is given by the rate of reactions (1)

$$\delta n_p = \delta n_e = -\delta n_n = \int_0^t (\mathrm{d}n_p/\mathrm{d}t)\mathrm{d}t. \tag{5}$$

The reaction rates of the two reactions have already been derived in ref. [7] wherein the angular integrals involved in the phase space integrals have been evaluated exactly. The results for the case of degenerate matter ($\mu_n, \mu_p, \mu_e \gg T$) are

$$\Gamma_1 = \frac{G^2 \cos^2\theta_c}{4\pi^5} T^3 \int_0^\infty \mathrm{d}x \, \frac{(x-\alpha)^2 + \pi^2}{1 + e^{x-\alpha}} I(\mu_n, Tx, \mu_p, \mu_e) \tag{6}$$

and

$$\Gamma_2 = \frac{G^2 \cos^2\theta_c}{4\pi^5} T^3 \int_0^\infty \mathrm{d}x \, \frac{(x+\alpha)^2 + \pi^2}{1 + e^{x+\alpha}} J(\mu_n, Tx, \mu_p, \mu_e), \tag{7}$$

where G is the Fermi constant, θ_c the Cabibbo angle, and $\alpha = \delta\mu/T$. In these equations I and J are complicated expressions given by (for detials of steps leading to these expressions see ref. [7])

$$\begin{aligned} I(E_1, E_2, E_3, E_4) = {} & k_-^{12} k_+^{34} I_1(E_1, E_2, E_3, E_4) + \tfrac{1}{6}(k_+^{34} - k_-^{12}) \\ & \times I_2(E_1, E_2, E_3, E_4) - \tfrac{1}{20} I_3(E_1, E_2, E_3, E_4), \end{aligned} \tag{8}$$

where

$$k_\pm^{ij} = E_i E_j \pm \tfrac{1}{2}(p_i^2 + p_j^2), \tag{9}$$

$$\begin{aligned} I_n(E_i, E_j, E_k, E_l) = {} & (P_{ij}^{2n-1} - p_{ij}^{2n-1})\theta(P_{kl} - P_{ij})\theta(p_{ij} - p_{kl}) \\ & + (P_{ij}^{2n-1} - p_{kl}^{2n-1})\theta(P_{kl} - P_{ij})\theta(p_{kl} - p_{ij})\theta(P_{ij} - p_{kl}) \\ & + (P_{kl}^{2n-1} - p_{kl}^{2n-1})\theta(P_{ij} - P_{kl})\theta(p_{kl} - p_{ij}) \\ & + (P_{kl}^{2n-1} - p_{ij}^{2n-1})\theta(P_{ij} - P_{kl})\theta(p_{ij} - p_{kl})\theta(P_{kl} - p_{ij}), \end{aligned} \tag{10}$$

$$P_i = (E_i, \mathbf{p}_i), \quad |\mathbf{p}_i| = p_i. \tag{12}$$

The net rate for the production of protons is therefore

$$\frac{dx_p}{dt} = \Gamma_1 - \Gamma_2 = A f(\delta\mu), \tag{13}$$

where

$$A = \frac{G^2 \cos^2\theta_c}{4\pi^5} T^3 \tag{14}$$

and

$$f(\delta\mu) = \int_0^\infty dx \left[\frac{(x-\alpha)^2 + \pi^2}{1 + e^{x-\alpha}} I - \frac{(x+\alpha)^2 + \pi^2}{1 + e^{x+\alpha}} J \right]. \tag{15}$$

For calculating energy dissipation according to eq. (3), only last three terms in eq. (4) contribute. Now

$$\frac{\partial P}{\partial n_i} = -\frac{\partial \mu_i}{\partial v} \tag{16}$$

and

$$\delta\mu(t) = \delta\mu(0) + \left(\frac{\partial(\delta\mu)}{\partial v}\right)_0 \delta v + \left(\frac{\partial \delta\mu}{\partial n_n}\right)_0 \delta n_n + \left(\frac{\partial \delta\mu}{\partial n_p}\right) \delta n_p + \left(\frac{\partial \delta\mu}{\partial n_e}\right) \delta n_e. \tag{17}$$

This leads to

$$\frac{d\delta\mu(t)}{dt} = \omega B \left(\frac{\Delta v}{v_0}\right) \cos\omega t - CAf(\delta\mu), \tag{18}$$

where

$$B = v_0 \left(\frac{\partial \delta\mu}{\partial v}\right)_0; \quad C = -v_0 \left(\frac{\partial \delta\mu}{\partial n_p} + \frac{\partial \delta\mu}{\partial n_e} - \frac{\partial \delta\mu}{\partial n_n}\right)_0 \tag{19}$$

and

$$w = \frac{2\pi}{\tau}. \tag{20}$$

Defining the bulk viscosity ζ by

$$\zeta = 2 \frac{(dw/dt)_{av}}{v_0} \left(\frac{v_0}{\Delta v}\right)^2 \frac{1}{\omega^2} \tag{21}$$

and using eqs (3), (4) and (16)–(18), we obtain

$$\zeta = \frac{1}{\pi\omega} \left(\frac{v_0}{\Delta v}\right) \int_0^\tau A f(\delta\mu) \sin(\omega t) dt. \tag{22}$$

For the description of dense nuclear matter we have used Walecka's mean field theory [16–18]. In the mean field approximation we allow the meson fields σ, ω and ρ to acquire

density dependent average values. From the effective Lagrangian one can then read off the effective masses and chemical potentials $\bar{m}_i$ and $\bar{\mu}_i$. The fermi momenta are, in turn, related to these effective quantities by

$$\bar{k}_i = (\bar{\mu}_i^2 - \bar{m}_i^2)^{1/2}. \tag{23}$$

The various chemical potentials and masses required in these expressions are obtained from the self-consistent solution of the mean field equations:

$$m_\sigma^2\bar{\sigma} + bm_n g_{\sigma N}^3\bar{\sigma}^2 + cg_{\sigma N}^4\bar{\sigma}^3 - g_{\sigma N}(n_p^s + n_n^s) = 0, \tag{24}$$

$$m_\omega^2\bar{\omega}_0 = g_{\omega N}(n_p + n_n), \tag{25}$$

$$m_\rho^2\bar{\rho}_0 = \tfrac{1}{2}g_{\rho N}(n_p - n_n), \tag{26}$$

$$m_n - \bar{m}_n = m_p - \bar{m}_p = g_{\sigma N}\bar{\sigma}_0, \tag{27}$$

$$\bar{\mu}_n = \mu_n - g_{\omega N}\bar{\omega}_0 + \tfrac{1}{2}g_{\rho N}\bar{\rho}_0, \tag{28}$$

$$\bar{\mu}_p = \mu_p - g_{\omega N}\bar{\omega}_0 - \tfrac{1}{2}g_{\rho N}\bar{\rho}_0. \tag{29}$$

The charge densities n_i and the scalar densities, n_i^s (see ref. [17]) are given in terms of these quantities by the expressions

$$n_i = 2\int \frac{\mathrm{d}^3p}{(2\pi)^3}\,[1 + \mathrm{e}^{(E_i - \bar{\mu}_i)/T}]^{-1}, \tag{30}$$

$$n_i^s = 2\bar{m}_i\int \frac{\mathrm{d}^3p}{(2\pi)^3}\frac{1}{E_i}\,[1 + \mathrm{e}^{(E_i - \bar{\mu}_i)/T}]^{-1}, \tag{31}$$

where

$$E_i = (p^2 + \bar{m}_i^2)^{1/2}. \tag{32}$$

Since the chemical potentials μ_n and μ_p are of the order of a few hundred MeV, even for temperatures of up to a few MeV the matter is completely degenerate leading to the following expansions (up to T^2 terms):

$$n_i = \frac{1}{3\pi^2}\left\{(\bar{\mu}_i^2 - \bar{m}_i^2)^{3/2} + \frac{\pi^2T^2}{2}\frac{(2\bar{\mu}_i^2 - \bar{m}_i^2)}{(\bar{\mu}_i^2 - \bar{m}_i^2)^{1/2}}\right\} \tag{33}$$

and

$$n_i^s = \frac{\bar{m}_i}{2\pi^2}\left\{(\bar{\mu}_i^2 - \bar{m}_i^2)^{1/2}\bar{\mu}_i - \bar{m}_i^2\cdot\ln\left(\frac{\bar{\mu}_i + \sqrt{\bar{\mu}_i^2 - \bar{m}_i^2}}{\bar{m}_i}\right) + \frac{\pi^2}{3}\frac{T^2\bar{\mu}_i}{(\bar{\mu}_i^2 - \bar{m}_i^2)^{1/2}}\right\}, \tag{34}$$

where $i = n, p$. To evaluate n_e, set $\bar{\mu}_e = \mu_e$ and $\bar{m}_i = m_e$ in (33). Using these expansions in (19) and going through some straightforward but tedious algebra one obtains

$$B = \frac{1}{3}\left(\frac{\bar{m}_n^2}{\bar{\mu}_n} - \frac{\bar{m}_p^2}{\bar{\mu}_p} - \frac{m_e^2}{\mu_e}\right) + \frac{2\pi^2}{9}T^2\left(\frac{\bar{\mu}_p}{\bar{\mu}_p^2 - \bar{m}_p^2} + \frac{\mu_e}{\mu_e^2 - m_e^2} - \frac{\bar{\mu}_n}{\bar{\mu}_n^2 - \bar{m}_n^2}\right)$$
$$- \frac{2}{3}\left(\frac{g_{\rho N}}{m_\rho}\right)^2(n_p - n_n) \tag{35}$$

and

$$C = \frac{1}{3}\left(\frac{\bar{k}_p^2}{\bar{\mu}_p n_p} + \frac{k_e^2}{\mu_e n_e} + \frac{\bar{k}_n^2}{\bar{\mu}_n n_n}\right) + \frac{2\pi^2 T^2}{9}\left(\frac{\bar{\mu}_p}{n_p \bar{k}_p^2} + \frac{\mu_e}{n_e k_e^2} + \frac{\bar{\mu}_n}{n_n \bar{k}_n^2}\right) + \left(\frac{g_{\rho N}}{m_\rho}\right)^2. \tag{36}$$

We have used the values of the various couplings $g_{\sigma n}, g_{\omega n}, g_{\rho n}, b$ and c fixed by Ellis *et al* [18], namely

$$g_{\sigma N}/m_\sigma = 0.0152502\,\text{MeV}^{-1},$$
$$g_{\omega N}/m_\omega = 0.011\,\text{MeV}^{-1},$$
$$g_{\rho N}/m_\rho = 0.011\,\text{MeV}^{-1},$$
$$b = 3.418 \times 10^{-3},$$
$$c = 0.0146.$$

With this set of parameters, one obtains [18]

$$n_B(\text{saturation}) = 0.153\,\text{fm}^{-3},$$
$$\text{binding energy}(E/n_B - m_N) = -16.3\,\text{MeV},$$
$$\text{charge symmetry energy} = 32.5\,\text{MeV},$$
$$\text{bulk modulus} = 300\,\text{MeV},$$
$$\text{Landau mass} = 0.83 m_N.$$

To obtain bulk viscosity, we solve the differential equation (18) for $\delta\mu(t)$, obtain $f(\delta\mu)$ by evaluating the integral in (15) and finally evaluate the integral in (21). These steps of course have to be done numerically. In the linear approximation which is valid for small perturbations, (21) can be solved analytically and we obtain

$$\zeta = \frac{A'B^2}{(A'C)^2 + \omega^2}\left[1 - \frac{\omega A'C}{\pi} \cdot \frac{1 - e^{-A'C\tau}}{(A'C)^2 + \omega^2}\right], \tag{37}$$

where

$$A' = \frac{A}{T}\int_0^\infty dx(I+J)\left[\frac{(x^2+\pi^2)e^x}{(1+e^x)^2} - \frac{2x}{1+e^x}\right]. \tag{38}$$

This integral of course still has to be evaluated numerically.

3. Results and discussion

The results for the nuclear matter viscosity as derived from eq. (21) are given in figures 1–3 for various values of temperatures, nuclear densities n_B, and periods of harmonic perturbations, τ. In each figure we have plotted the viscosity as a function of relative volume perturbation $(\Delta v/v_0)$ for different temperatures. These figures correspond to $n_B = 0.3\,\text{fm}^{-3}$ and $0.6\,\text{fm}^{-3}$ (figures 1a and 1b respectively) and $\tau = 0.0005$ s. In

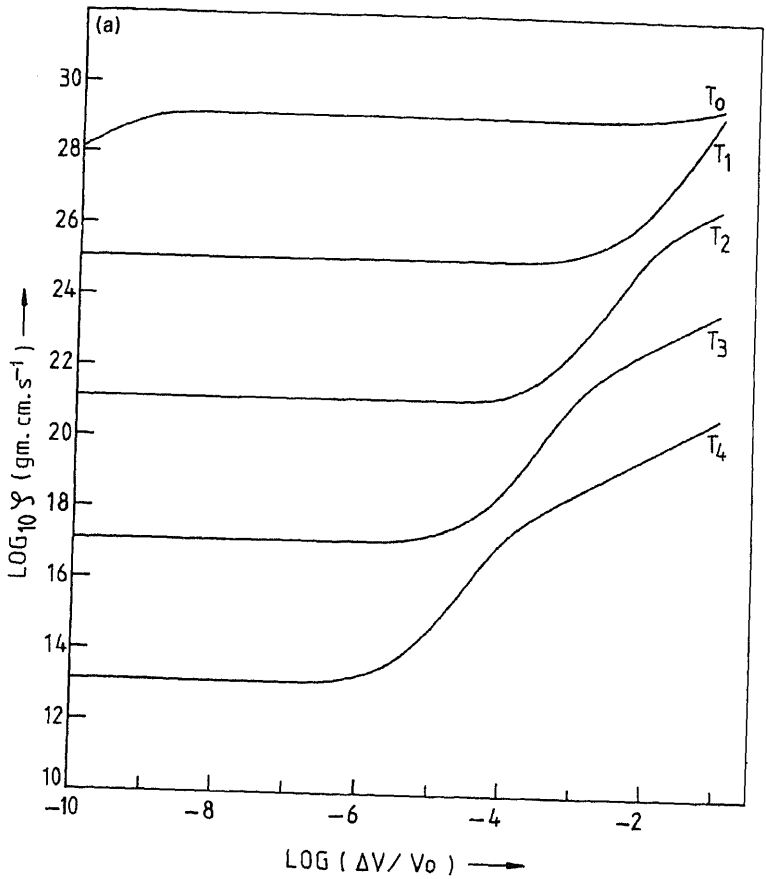


Figure 1a. Log of viscosity ζ vs log of amplitude of perturbation $(\Delta v/v_0)$ for $n_B = 0.3\,\mathrm{fm}^{-3}, \tau = 0.0005\,\mathrm{s}$. The five curves T_0, T_1, T_2, T_3, T_4 correspond to temperatures $T = 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\,\mathrm{MeV}$ respectively.

igure 2 we show our results for three different time periods, viz., $\tau_1 = 0.001\,\mathrm{s}$, $_2 = 0.0005\,\mathrm{s}$ and $\tau_3 = 0.0001\,\mathrm{s}$ and two different temperatures, viz., 1 MeV and .001 MeV. We find that by and large the dependence of viscosity on the perturbation mplitude is similar to that obtained by Madesen [12] for strange quark matter. For small erturbations the viscosity is nearly independent of the amplitude of perturbation whereas or large perturbations there is a sharp increase in the viscosity by many orders of agnitude. As far as dependence on temperature is concerned, Madesen [12] found that ie viscosity increases with increasing temperature initially but for $T > 0.1\,\mathrm{MeV}$ it starts ecreasing. A similar behaviour was reported by Goyal *et al* [13] in a more detailed alculation of the strange matter viscosity. In the present calculation we find that the uclear matter viscosity continues to increase with temperature at least up to 2 MeV.

We now compare our results with those of Haensel and Schaeffer [10]. In the erturbation independent regime the viscosity is nearly proportional to the fourth power f temperature in agreement with the results of URCA processes as given by Haensel and chaeffer [10]. However, at higher amplitudes of perturbation, the behaviour of viscosity

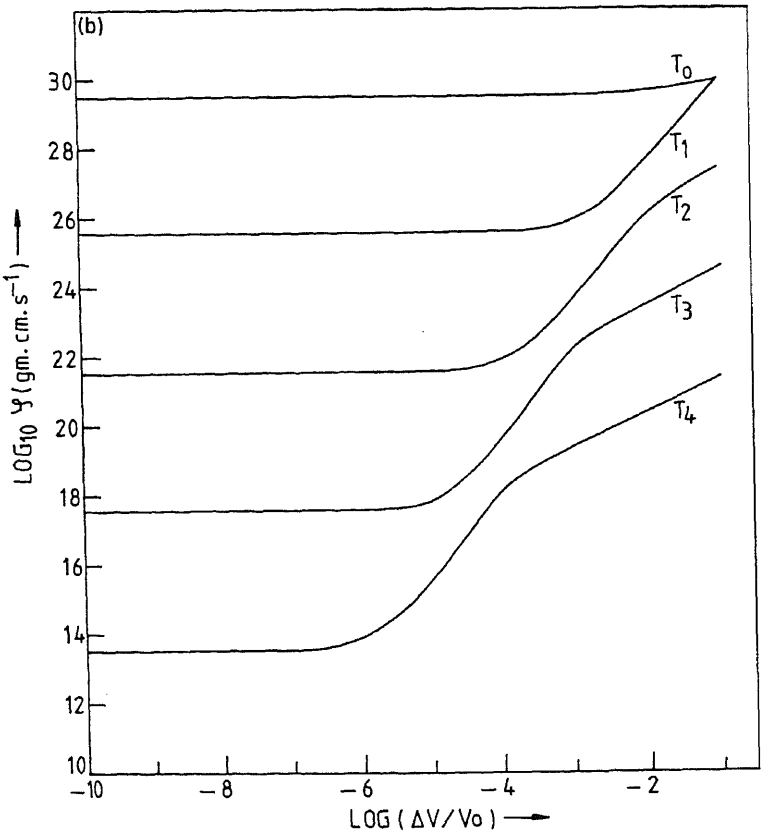


Figure 1b. Same as figure 1a for $n_B = 0.6\,\mathrm{fm}^{-3}$, $\tau = 0.0005$ s. The five curves correspond to the same five temperatures.

becomes markedly different. It increases very rapidly with $\Delta v/v_o$ except at the highest temperatures considered ($T \gtrsim 1$ MeV). This clearly indicates that the linear approximation is no longer valid at temperatures $\lesssim 0.1$ MeV. The lower the temperature, the smaller is the region of validity of linear approximation. Thus for $T = 10^{-4}$ MeV the linear approximation is valid only up to $\Delta v/v_0 \lesssim 10^{-6}$, whereas for $T = 0.1$ MeV, it is valid up to $\Delta v/v_0 \lesssim 10^{-3}$.

At higher amplitudes, even the temperature dependence becomes quite different and more involved than the simple T^4 behaviour expected in the linear range. For $\Delta v/v_0 \gtrsim 10^{-3}$ the viscosity increases much less rapidly than what is dictated by the T^4 law. As T increases by four orders (from 10^{-4} MeV to 1 MeV), viscosity increases by only 10–11 orders of magnitude instead of sixteen.

From figure 2 we find that even the τ^2 behaviour of viscosity is valid only at high temperatures. At lower temperatures ($T = 10^{-4}$ MeV), increase of τ from 0.0001 s to 0.001 s increases viscosity by six orders instead of just two. This is so even for small perturbations. As can be seen from (37) which is valid for small perturbations, the frequency ($\omega = 2\pi/\tau$) dependence of viscosity is far more complicated than the simple relation, $\rho \propto \omega^{-2}$ obtained by Haensel and Schaeffer [10].

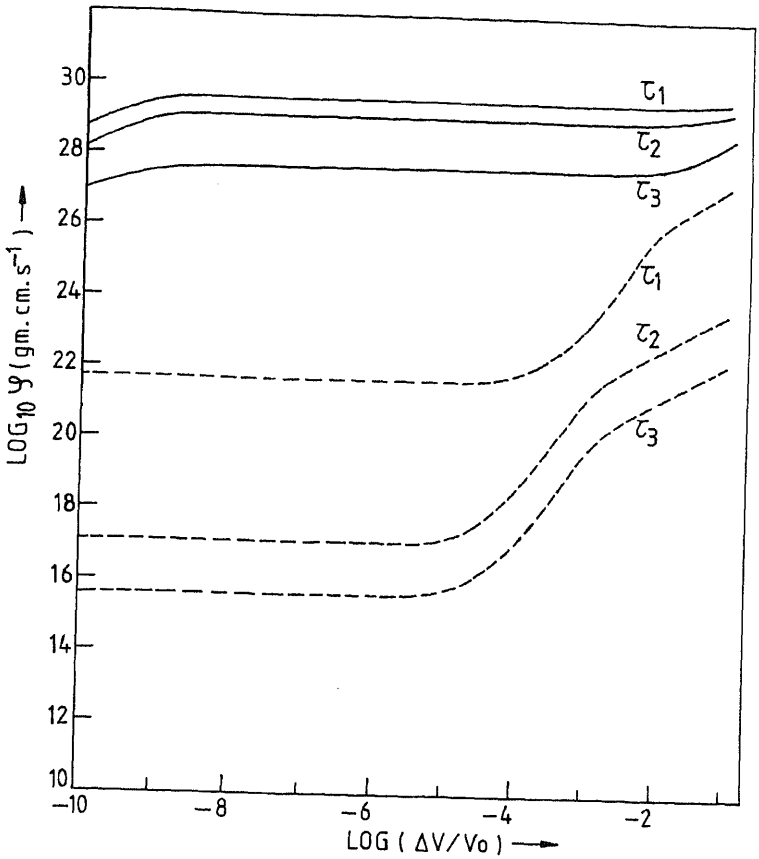


Figure 2. Log of viscosity ζ vs log of amplitude of perturbation $(\Delta v/v_0)$ for three time periods $\tau_1 = 0.001$ s, $\tau_2 = 0.0005$ s and $\tau_3 = 0.0001$ s, and two different temperatures, $T = 1$ MeV (solid curves) and $T = 0.001$ MeV (dashed curves). $n_B = 0.3$ fm^{-3} for all the sets.

The dependence of viscosity on nuclear density is rather weak (figure 3), and is adequately described by the $\rho \propto n^{1/3}$ relation given by Haensel and Schaeffer. As for the dependence on proton fraction $x(= n_p/n_B)$, in our calculation it cannot be judged as x appears as a parameter in our expressions only through its dependence on n_B. As n_B changes from n_0 to $10n_0$, x increases by a factor of five (figure 3). Thus the variation in the viscosity with the nuclear density n_B includes the implied dependence on x as well.

One of the main applications of the bulk viscosity is in determining the damping time of star vibrations. A rough estimate of the damping time can be obtained in the manner of Sawyer [11] who has obtained the following result:

$$\tau_d = \frac{1}{60}\left(\frac{\delta\rho}{\rho_0}\right)^2 \rho_m \omega^2 R^2 \frac{2}{\zeta(\delta\rho/\rho_0)\omega^2} = R^2\zeta^{-1}$$
$$= \frac{1}{30}\rho_m R^2 \zeta^{-1}. \tag{39}$$

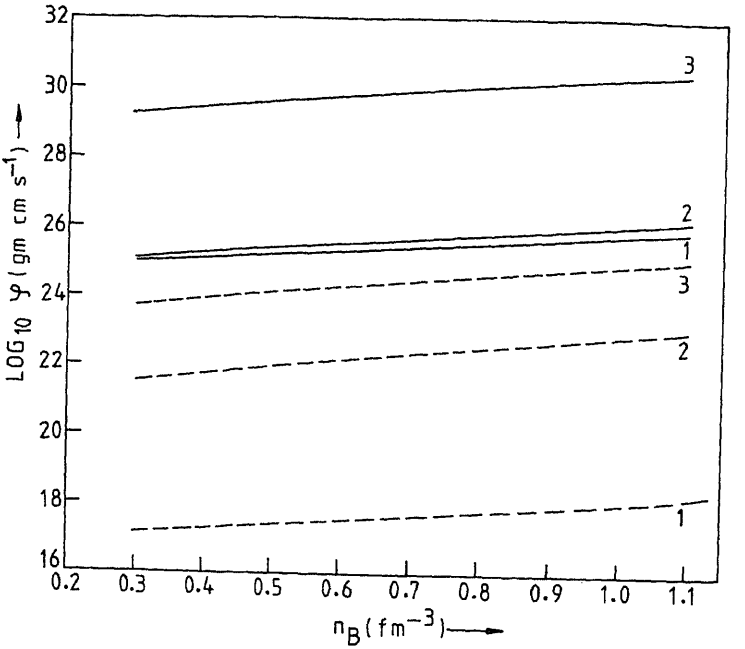


Figure 3. Log of viscosity ζ vs the nuclear density n_B. Solid curves correspond to $T = 0.1$ MeV and dashed curves to $T = 0.001$ MeV. Curves labeled 1, 2, 3 correspond respectively to $(\Delta v/v_0) = 10^{-5}, 10^{-3}, 10^{-1}$ for both temperatures. $\tau = 0.0005$ s for all the sets.

For a typical neutron star with mass $1.4 M_\theta, \rho_m = 4 \times 10^{14}$g/cc and radius $R \sim 10$ km, $\tau_d \approx 1.5 \times 10^{25} \zeta^{-1}$ s. It is clear that since the viscosity increases by many orders of magnitude as the amplitude of perturbation increases, particularly for low temperatures, the high amplitude perturbations will die out very rapidly. For low amplitude perturbations our results for viscosity are somewhat larger (by about an order of magnitude) than those obtained in [10], thus yielding considerably shorter damping times.

A comparison with the results for the strange quark matter is not so straightforward. In our case the viscosity increases with temperature whereas in the case of strange stars viscosity exhibits a more complicated behaviour: it first increases with temperature but for $T > 0.1$ MeV it decreases with further increase of temperature [12, 13]. Also as a function of n_B, the viscosity exhibits the same peculiar behaviour i.e. increases with n_B for low temperatures (< 0.1 MeV), but decreases with increasing n_B for higher temperatures (> 0.1 MeV) [13]. However, generally for n_B around $3n_0$ and for a temperature around 0.1 MeV, the strange star viscosity is about an order of magnitude lower than our result [13].

As compared to the modified URCA process which has a T^6 dependence on temperature, for the URCA process it is close to T^4. At a temperature close to 0.1 MeV, our calculation yields a value of bulk viscosity of 9 to 10 orders of magnitude more than for the modified URCA process as obtained by Sawyer [9].

dramatic difference between the time scales of cooling of neutron star cores with
$_t$ and $x < x_{crit}$ cases have already been highlighted in ref. [7]. Since our emissivity
] are the same as those obtained by Lattimer *et al* [6] within a factor of two or so,
ıe scales for cooling are also nearly the same. However our larger values of
ty imply a higher stability for the rapidly rotating star immediately after it is born.

nces

Finzi and R A Wolf, *Astrophys. J.* **153**, 835 (1968)
Cutler, L Lindblom and R J Splinter, *Astrophys. J.* **363**, 603 (1990)
Cutler and L Lindblom, *Astrophys. J.* **314**, 234 (1987)
Witten, *Phys. Rev.* **D30**, 272 (1984)
Farhi and R L Jaffe, *Phys. Rev.* **D30**, 2379 (1984)
M Lattimer, C J Pethick, M Prakash and P Haensel, *Phys. Rev. Lett.* **66**, 2701 (1991)
Wadhwa, V K Gupta, S Singh and J D Anand, *J. Phys.* **G9**, 1137 (1995)
D Wang and T Lu, *Phys. Lett.* **B148**, 211 (1984)
ymond F Sawyer, *Phys. Rev.* **D39**, 3804 (1989)
Haensel and R Schaeffer, *Phys. Rev.* **D45**, 4708 (1992)
F Sawyer, *Phys. Lett.* **B233**, 412 (1989)
s Madesen, *Phys. Rev.* **D46**, 3290 (1992)
Goyal, V K Gupta, Pragya and J D Anand, *Z. Phys.* **A349**, 93 (1994)
e, for example, N Iwamoto, *Ann. Phys. NY* **141**, 1 (1982)
Wadhwa, J D Anand and V K Gupta, *Pramana – J. Phys.* **45**, 195 (1995)
D Walecka, *Ann. Phys. NY* **83**, 491 (1974)
D Serot and J D Walecka, in *Advances in nuclear physics* **16**, 1 (1986)
Ellis, J I Kapusta and K A Olive, *Nucl. Phys.* **B348**, 345 (1991)

physics

November 1997
pp. 455–460

Application of density-functional perturbation theory to calculate nonlinear polarizabilities of helium-like systems

ARUP BANERJEE and MANOJ K HARBOLA
Laser Programme, Centre for Advanced Technology, Indore 452 013, India
Email: banerjee@cat.ernet.in

MS received 29 April 1997; revised 16 August 1997

Abstract. Density-based perturbation theory within the Hohenberg–Kohn (HK) formalism of density functional theory (DFT), developed recently by us, is employed to calculate hyperpolarizabilities of helium-like ions from their ground-state densities obtained from their respective Hylleraas wavefunctions. The only approximation made is that of the local density (LDA) for exchange and correlation. Use of densities – instead of wavefunctions – in density-based perturbation theory together with simple approximate energy functionals makes our calcualtions much simpler than those based on wavefunctions. They lead, however, to accurate results.

Keywords. Static polarizabilities; perturbation theory; density functional theory.

PACS Nos 31.15; 32.10

Recently we have developed the density-based perturbation theory (DFPT) [1] within the Hohenberg–Kohn (HK) formalism of density functional theory (DFT). In this theory perturbative calculations are performed in terms of the density. The theory is different from the widely used Kohn–Sham version [2] of density-functional theory where an orbital based approach is used. The two key theorems of DFPT are: (a) the $(2n+1)$ theorem which states that the density correct to order 'n' in terms of perturbation parameter is sufficient to give energy up to order $(2n+1)$, and (b) as a corollary, if the density is known exactly to order $(n-1)$, the even order energy change $E^{(2n)}$ is variational with respect to the nth order induced density $\rho^{(n)}$. Of these, the $(2n+1)$ theorem has been mentioned [3] before in the literature. However, it is only now that perturbation theory in terms of density and associated theorems have been fully developed. This paves the way for obtaining both linear as well as non-linear response properties of a system directly from its density. In this paper we apply the $(2n+1)$ theorem and its even-order corollary in DFPT to calculate variationally the electric dipole polarizabilities α and hyperpolarizabilities γ for two-electron helium-like ions from their ground-state densities. Thus α is obtained by minimizing the second-order energy change $\Delta E^{(2)}$ and using the relationship

$$\Delta E^{(2)} = -\tfrac{1}{2}\alpha\mathcal{E}^2. \tag{1}$$

versions (for spherically symmetric systems considered in this paper) of more general equations in which α and γ are tensors. Our motivation in carrying out the calculation is to demonstrate that with approximate exchange-correlation energy functionals, density-based DFPT can be applied to calculate linear and non-linear response properties of these systems. We note that the density-based calculation for linear response properties of atoms have been performed earlier [3,4]. However, it is only after the development of density-based variation-perturbation theory within the Hohenberg–Kohn (HK) formalism of DFT that, the calculation of nonlinear response properties of electronic systems from their ground-state densities has become possible. Further, since the energy changes are calculated from the ground-state densities, rather than the wavefunctions of the atoms, the present method is much simpler and numerically easier than its wavefunctional counterpart.

In the present paper we employ the near exact densities constructed [5] from explicitly correlated two-electron Hylleraas wavefunctions [6]. The approximate forms for the kinetic and exchange-correlation energy functionals employed are described later in the paper.

In the following we briefly describe DFT [2] and its perturbative extension along with the energy functionals chosen for our calculations. We then present our calculations and results and follow them with a discussion.

The basic variable in DFT is the ground-state density $\rho(\mathbf{r})$ and all the properties of the system are expressed as functionals of it. In particular, the ground-state energy of a system is also written as the functional

$$E[\rho] = F[\rho] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{3}$$

where $F[\rho]$ is a universal functional of $\rho(\mathbf{r})$ and comprises the kinetic and the interaction energy of the electrons, and $v_{\text{ext}}(\mathbf{r})$ is the external potential including $v^{(1)}(\mathbf{r}) = \mathcal{E}r\cos\theta$ due to the applied field $\mathcal{E}$. The functional $F[\rho]$ is generally split as

$$F[\rho] = T_s[\rho] + E_H[\rho] + E_{\text{xc}}[\rho], \tag{4}$$

where $T_s[\rho]$ is the kinetic energy of the non-interacting electrons of the same density as the interacting ones,

$$E_H[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \tag{5}$$

is the Hartree energy, and $E_{\text{xc}}[\rho]$ is the exchange-correlation energy. The difference between the true kinetic energy of the electrons and $T_s[\rho]$ is absorbed into the exchange-correlation energy $E_{\text{xc}}[\rho]$. Employing the variational principle for the energy, both the ground-state density and energy are obtained by minimizing the energy functional with respect to the density. This leads to the following Euler equation for the density

$$\frac{\delta E[\rho]}{\delta\rho(\mathbf{r})} - \mu = 0, \tag{6}$$

where μ is the Lagrange multiplier to ensure that the number of electrons is conserved. When the system is subjected to an external applied field, both the ground-state density and energy change from unperturbed ρ_0 and E_0 to $\rho_0 + \Delta\rho$ and $E_0 + \Delta E$ respectively. In DFPT these changes are expanded in perturbation series as

$$\Delta\rho = \sum_i \lambda^i \rho^{(i)}, \tag{7}$$

with

$$\int \rho^{(i)}(\mathbf{r})\,\mathrm{d}\mathbf{r} = 0 \tag{8}$$

for each order, and

$$\Delta E = \sum_i \lambda^i E^{(i)}, \tag{9}$$

where λ is the perturbation parameter. The energy changes to different orders are calculated in terms of $\rho^{(i)}$. By the $(2n+1)$ theorem of DFPT, energy up to $E^{(2n+1)}$ requires knowledge of ρ only to order n. As such [1]

$$E^{(2)} = \int v^{(1)}(\mathbf{r})\rho^{(1)}(\mathbf{r})\mathrm{d}\mathbf{r} + \frac{1}{2}\int \frac{\delta^{(2)}F[\rho_0]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)}\,\rho^{(1)}(\mathbf{r}_1)\rho^{(1)}(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2, \tag{10}$$

involves density up to $\rho^{(1)}$, and

$$\begin{aligned} E^{(4)} = {} & \frac{1}{2}\int \frac{\delta^{(2)}F[\rho_0]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)}\rho^{(2)}(\mathbf{r}_1)\rho^{(2)}(\mathbf{r}_2)\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \\ & + \frac{1}{2}\int \frac{\delta^{(3)}F[\rho_0]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)\delta\rho(\mathbf{r}_3)}\rho^{(1)}(\mathbf{r}_1)\rho^{(1)}(\mathbf{r}_2)\rho^{(2)}(\mathbf{r}_3)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2\mathrm{d}\mathbf{r}_3 \\ & + \frac{1}{24}\int \frac{\delta^{(4)}F[\rho_0]}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)\delta\rho(\mathbf{r}_3)\delta\rho(\mathbf{r}_4)} \\ & \times \rho^{(1)}(\mathbf{r}_1)\rho^{(1)}(\mathbf{r}_2)\rho^{(1)}(\mathbf{r}_3)\rho^{(1)}(\mathbf{r}_4)\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3\,\mathrm{d}\mathbf{r}_4 \end{aligned} \tag{11}$$

is given by density up to order 2. (For the present problem all the odd orders of energy vanish due to spherical symmetry). All the functional derivatives in the expressions above have been taken at the ground-state density ρ_0. Furthermore, by the even-order corollary of the theorem, $E^{(2)}$ is variational with respect to $\rho^{(1)}$ and, with known $\rho^{(1)}$, $E^{(4)}$ is variational with respect to $\rho^{(2)}$. This results in the equations for the induced densities $\rho^{(1)}$ and $\rho^{(2)}$ which can be solved to obtain these quantities. The induced densities and the energies can also be determined by applying the variational procedure sequentially [7] in two steps. First $\rho^{(1)}$ is determined by minimizing $E^{(2)}$. Then $\rho^{(1)}$ is kept fixed and $E^{(4)}$ is minimized to determine $\rho^{(2)}$. It is this variational approach that we take in this paper. Notice that the expression for $E^{(4)}$ does not involve explicitly the applied external field. Thus the information about the field is carried to $E^{(4)}$ by the induced density $\rho^{(1)}$.

Before we present our results, it is worth pointing out the difference between the density-based perturbation theory used in this paper and the Kohn–Sham perturbation theory. In the density-based approach the correction from $E^{(3)}$ onwards are given by the

universal functional $F[\rho]$ alone. On the other hand the Kohn–Sham approach employs the full Hamiltonian (including the applied perturbation) for the energy changes to all orders. In principle both should lead to the same result. However, from the application point of view the two theories differ because of their treatment of the kinetic energy. In Kohn–Sham theory the non-interacting component of the kinetic energy is treated *exactly* whereas in the density-based approach it has to be approximated for systems with number of electrons greater than two. Although many approximate functionals for the kinetic energy exists [2], it is not clear which one should lead to accurate response properties for a many-electron system; the application of Kohn–Sham approach [7, 8] has no such limitation.

To perform variational calculations, the question arises how to choose the functional form for the induced densities $\rho^{(1)}$ and $\rho^{(2)}$. For this purpose, we take a cue from the hydrogen atom problem which can be solved exactly [9, 10] and employ similar form for the variational densities. They are taken to be

$$\rho^{(1)}(\mathbf{r}) = \Delta_1(r)\cos\theta\rho_0(r) \tag{12}$$

and

$$\rho^{(2)}(\mathbf{r}) = [\Delta_2(r) + \Delta_3(r)\cos^2\theta]\rho_0(r) + \lambda\rho_0(\mathbf{r}), \tag{13}$$

where

$$\Delta_i(r) = a_i r + b_i r^2 + c_i r^3 + d_i r^4 + e_i r^5 + \ldots; \quad i = 1, 2, 3 \tag{14}$$

with $a_i, b_i, \ldots, e_i$ being the variational parameters which are optimized to obtain the minimum of energy. Notice that with these choices for the induced densities, eq. (8) is automatically satisfied for $\rho^{(1)}$; to satisfy it for $\rho^{(2)}$, λ is appropriately chosen for each set of parameters.

As stated in the beginning, density-based calculations for an N-electron system perforce employ approximate energy functionals, since the exact ones are not known. For the exchange and correlation energy we make the local-density approximation (LDA). Thus

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho] = E_{\mathrm{x}}^{\mathrm{LDA}}[\rho] + E_{\mathrm{c}}^{\mathrm{LDA}}[\rho], \tag{15}$$

where the exchange energy

$$E_{\mathrm{x}}^{\mathrm{LDA}}[\rho] = -\frac{3}{4}(3/\pi)^{1/3}\int \rho^{4/3}(\mathbf{r})\mathrm{d}\mathbf{r} \tag{16}$$

and the correlation energy [11]

$$E_{\mathrm{c}}^{\mathrm{LDA}}[\rho] = \int \rho(\mathbf{r})c\left[(1+x^3)\ln\left(1+\frac{1}{x}\right) + \frac{x}{2} - x^2 - \frac{1}{3}\right]\mathrm{d}\mathbf{r} \tag{17}$$

with $c = -0.0333$, $x = r_s/A$, $A = 11.4$, where $r_s = [3/4\pi\rho(\mathbf{r})]^{1/3}$. On the other hand, for two-electron systems the non-interacting kinetic energy is given exactly by the von-Weizsacker energy functional [12]

$$T_{\mathrm{W}} = \frac{1}{8}\int \frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}\,\mathrm{d}\mathbf{r}. \tag{18}$$

Table 1. Electric dipole polarizabilities α in atomic units for various helium-like ions obtained from their Hylleraas densities. Theoretical values from the other *ab-initio* schemes, and the experimental numbers are also given.

Atom/Ion	Present results	*Ab-initio*	Expt.[c]
H^-	231.23	206.39[a]	—
He	1.3891	1.3834[a]	1.3838
Li^+	0.1928	0.1925[a]	0.1911
Be^{2+}	0.0523	0.0519[b]	—
B^{3+}	0.0197	0.0196[c]	—

[a] Ref. [15], [b] Ref. [16] and [c] Ref. [17].

Table 2. Electric dipole hyperpolarizabilities in atomic units for various helium-like ions obtained from their Hylleraas densities. Theoretical values from the *ab-initio* schemes, and the experimental numbers are also given. Experimental values differ from each other because of the different methods used to determine them.

Atom/Ion	Present results	*Ab-initio*	Expt.
H^-	1.12×10^8	0.803×10^{8a}	—
He	47.16	43.6[b]	35.7[e] 41.7[f] 42.6[g] 52.4[h]
Li^+	0.2471	0.2427[c]	—
Be^{2+}	0.0087	0.0085[d]	—
B^{3+}	0.0007	0.0007[d]	—

[a] Ref. [18], [b] Ref. [19], [c] Ref. [14] and [17], [d] Ref. [17], [e] Ref. [24], [f] Ref. [25], [g] Ref. [26] and [27], [h] Ref. [20].

In tables 1 and 2 we show the polarizabilities and hyperpolarizabilities, respectively, for two-electron systems H^-, He, Li^+, Be^{2+} and B^{3+} obtained by us. Although this paper is devoted to the calculation of nonlinear response properties of two-electron systems, we nonetheless include the numbers obtained for linear polarizabilities also in order to provide a complete picture. Also shown are the theoretical results obtained by other *ab-initio* methods [13–19] along with available experimental results [20–27]. The ground-state density employed by us corresponds to 204-term Hylleraas type wavefunctions [5].

It is evident from the results that they are highly accurate in comparison to the numbers available by other methods. We also note that the numbers for polarizability are also consistent with those obtained previously from a similar calculation [4] but with a different form for the induced density $\rho^{(1)}$. The maximum deviation for both α and γ occurs for H^- for which the effect of Coulomb correlations is the largest and therefore the LDA is relatively less accurate. We wish to point out that we have performed our calculations with six parameters for α and fourteen parameters for γ. On the other hand, if calculations were to be performed using the wavefunctions directly, one would require 612 parameters for α and 1224 for γ – $3N$ [13] and $6N$ [14] parameters are needed for α and γ respectively, if an N term Hylleraas-type wavefunction is employed. We have also checked our results by increasing the number of parameters in $\Delta_i(r)$. They do not change appreciably which indicates their proper covergence.

As pointed out before, the calculations reported here differ from the widely used Kohn–Sham approach of density functional theory and its perturbative extensions [8, 28, 29]. The Kohn–Sham approach is an orbital based scheme, and as such its perturbation theory methodology [7, 29, 30] is similar to that of coupled Hartree–Fock theory [31]. On the other hand present calculations are performed entirely in terms of the density.

To conclude, we have shown that DFPT can be employed with approximate functionals for the exchange-correlation energies (LDA) to obtain reasonably accurate results for both linear as well as non-linear response properties of helium-like systems.

References

[1] M K Harbola and A Banerjee, *Phys. Lett.* **A222**, 315 (1996)
[2] R M Dreizler and E K U Gross, *Density functional theory* (Springer, Berlin, 1990)
[3] L J Bartolotti, *Proc. Indian Acad. Sci.* **106**, 103 (1994)
[4] M K Harbola, *Phys. Rev.* **A48**, 2696 (1993) *Chem. Phys. Lett.* **217**, 461 (1994); *Int. J. Quantum Chem.* **51**, 201 (1994)
[5] T Koga , Y Kasai and A J Thakkar, *Int. J. Quantum Chem.* **46**, 689 (1993)
[6] E A Hylleraas, *Z. Phys.* **48**, 469 (1928)
E A Hylleraas *Z. Phys.* **54**, 347 (1928)
[7] M K Harbola and A Banerjee, *Phys. Rev.* **A54**, 283 (1996)
[8] S K Ghosh and B M Deb, *Chem. Phys.* **71**, 295 (1982)
L J Bartolotti, *J. Chem. Phys.* **80**, 5687 (1984); *J. Phys. Chem.* **90**, 5518 (1986)
[9] C A Coulson, *Proc. R. Soc. Edinburgh* **A61**, 20 (1941)
[10] G S Sewell, *Proc. Cambridge Phys. Soc.* **45**, 679 (1949)
[11] O Gunnarsson and B I Lundquist, *Phys. Rev.* **B13**, 4274 (1979)
[12] C F von Weizsacker, *Z. Phys.* **96**, 431 (1935)
[13] K T Chung and R P Hurst, *Phys. Rev.* **152**, 35 (1966)
[14] M N Grasso, K T Chung and R P Hurst, *Phys. Rev.* **167**, 167 (1968)
[15] R M Glover and F Weinhold, *J. Chem. Phys.* **65**, 4913 (1976)
[16] W Muller, J Flesch and W Meyer, *J. Chem. Phys.* **80**, 3297 (1984)
[17] D M Bishop and M Rerat, *J. Chem. Phys.* **91**, 5489 (1990)
[18] J Pipin and D M Bishop, *J. Phys.* **B25**, 17 (1992)
[19] J E Rice, P R Taylor, T J Lee and J Almolf, *J. Chem. Phys.* **94**, 4972 (1991)
[20] A D Buckingham and D A Dunmur, *Trans. Faraday Soc.* **64**, 776 (1968)
[21] J N Wilson and R M Curtis, *J. Phys. Chem.* **74**, 187 (1970)
[22] J Leonard, *At. Data Nucl. Data Tables* **14**, 21 (1974)
[23] D Gugan and G W Michel, *Mol. Phys.* **39**, 783 (1980)
[24] M Lundeen, S Y Hou and Z W Nibler, *J. Chem. Phys.* **79**, 6301 (1983)
[25] H T Lehmeier, W Leupacher and A Penzkofer, *Opt. Comm.* **56**, 67 (1985)
[26] V Mizrahi and D P Shelton, *Phys. Rev. Lett.* **55**, 696 (1985)
[27] D P Shelton, *J. Chem. Phys.* **84**, 404 (1986)
[28] G D Mahan and K R Subbaswami, *Local density theory of polarizability*, (Plenum, New York, 1990)
[29] X Gonze, *Phys. Rev.* **A52**, 1096 (1995)
[30] X Gonze and J P Vigneron, *Phys. Rev.* **B39**, 13120 (1989)
[31] P W Langhoff, J D Lyons and R P Hurst, *Phys. Rev.* **148**, 18 (1966)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 5,
November 1997
pp. 461–478

Polarization and dichroism in atomic Auger spectroscopy following absorption of circularly polarized light

S SEN
Department of Physics and Meteorology, Indian Institute of Technology, Kharagpur 721 302, India
Email: sandphy@phy.iitkgp.ernet.in

MS received 11 April 1997; revised 7 October 1997

Abstract. Spin resolved Auger current is shown to be different for the absorption of left and right circularly polarized light. This current is produced in non-radiative spontaneous decay of photoexcited atoms or of excited atomic photoions. Circular dichroism in such polarized Auger electrons exists even in the absence of spin–orbit interaction. Physical and geometrical conditions, necessary for the occurrence of dichroic effects in atomic Auger spectroscopy, have been obtained. Calculations for Ba and Xe show that the effect is substantially large and can be readily observed.

Keywords. Spin-polarization; circular dichroism; orientation and alignment parameters; spin orbit interaction (SOI); Auger electron.

PACS No. 33.55

1. Introduction

In a recent paper [1], we have investigated the existence of dichroic effects in Auger electron spectroscopy (AES), following absorption of circularly polarized (CP) light in a free and unpolarized rotating linear molecule in electric dipole ($E1$) approximation. The photoexcited molecule AB^* or the excited photoion AB^{+*} in eq. (I.1) of [1] remembers the helicity of the absorbed photon which created it. This memory is reflected in its subsequent radiationless decay. In other words, it means that the spin-resolved Auger current produced after the absorption of left circularly polarized (LCP) radiation is different from that emitted when the absorbed electromagnetic wave is right circularly polarized (RCP).

In the next section of this paper we show that the conditions in which one can observe circular dichroism (CD) in Auger spectroscopy of atoms are almost identical to those [1] of linear molecules. We, in particular, find that in order to study this effect in atoms, it is necessary to analyse spin of Auger electrons also but without taking the spin-orbit interaction (SOI) into account. Dichroism may exist both in the spin-resolved differential as well as integrated Auger currents. It does not necessarily vanish in any particular experimental configuration, including those which correspond to achiral geometries [2]. However, the dichroic effect will be absent if the integrated Auger current is spin-resolved perpendicular to the direction of incidence of CP ionizing radiation or Auger electrons are polarized perpendicular to the plane which contains both departing

electrons as well as incident photon beam. (These two and other experimental geometries are explained below in figure 1.) Another important feature is that the CD in Auger spectroscopy both of atoms and of rotating linear molecules is described by the same set of equations containing up to three parameters. But the dynamical factors to be used for the former are, of course, naturally different from those needed for the latter targets.

In order to see the magnitude of dichroic effects in atomic AES, we have performed in § 3 of this paper numerical calculations for Ba and Xe atoms. Here, the Auger currents are produced following photoionization in $5p$ and $4d$ subshells of Ba and Xe, respectively, after the absorption of a CP photon. In that section we have also analysed the spin polarization of Auger electrons emitted from these two atoms in the absence of SOI. The conclusion of this work is presented in § 4.

2. Theory for circular dichroism in atomic Auger spectroscopy

The processes of interest to us in the present paper are

$$h\nu_r + X(L_0S_0J_0M_0) \rightarrow X^*(LSJM) \tag{1a}$$

followed by

$$X^*(LSJM) \rightarrow X^+(L_fS_fJ_fM_f) + e_a(\mathbf{k}_a;\ \hat{u}_a\mu_a) \tag{1b}$$

and

$$h\nu_r + X(L_0S_0J_0M_0) \rightarrow X^{+*}(LSJM) + e_p \tag{2a}$$

followed by

$$X^{+*}(LSJM) \rightarrow X^{++}(L_fS_fJ_fM_f) + e_a(\mathbf{k}_a;\ \hat{u}_a\mu_a) \tag{2b}$$

caused by the absorption of a photon of frequency ν_r in a free and unpolarized atom X. While X^* in (1) is the photoexcited neutral atom, X^{+*} in (2) is the excited photoion formed after the emission of photoelectron e_p which remains unobserved. Emission of the Auger electron e_a results in the formation of residual singly and doubly charged atomic ions X^+ and X^{++} in (1b) and (2b), respectively. (L_0, L, L_f) and (S_0, S, S_f) are the respective total orbital and spin angular momenta of (X, X^*, X^+) in (1) and of (X, X^{+*}, X^{++}) in (2). We also have in these equations the total angular momenta $\mathbf{J}_0 = \mathbf{L}_0 + \mathbf{S}_0$, $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and $\mathbf{J}_f = \mathbf{L}_f + \mathbf{S}_f$ with their respective projections $M_0\hbar$, $M\hbar$ and $M_f\hbar$ along the polar axis of the space-frame. Further in (1b) and (2b), $\mathbf{k}_a(k_a, \hat{k}_a)$ is the propagation vector of Auger electron moving with kinetic energy $\epsilon_a = \hbar^2k_a^2/2m$ and the component $\mu_a\hbar(\mu_a = \pm 1/2)$ of its spin angular momentum along the spin quantization direction $\hat{u}_a(\theta_a, \phi_a)$ defined in the photon-frame coordinate system OXYZ in figure 1(a).

Spin polarization in the presence of spin-dependent interactions of Auger electrons emitted in the non-radiative decay of an inner atomic shell vacancy created in the excitation or ionization of unpolarized atoms by particle impact or photon absorption has theoretically been considered by many workers [3–7]. The atomic expressions derived in [3–7] can readily be written in the form of eq. (II.20). The Auger decay matrix elements

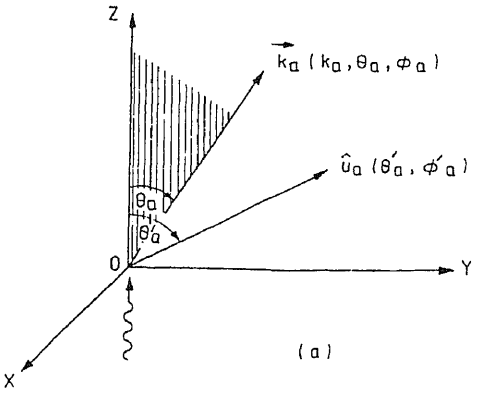


Figure 1(a). General geometry of an experiment used to study spin-polarization of Auger electrons. OXYZ is the photon (or laboratory) frame of reference. Its polar (OZ)-axis is along the direction of incidence of circularly polarized or unpolarized photon absorbed before Auger emissions. (If incident radiation is linearly polarized then its electric field vector is in the direction of OZ). Reaction (or scattering) plane contains the OZ-axis and the propagation vector $\mathbf{k}_a(k_a, \theta_a, \phi_a)$ of Auger electrons. Spin of Auger electrons is quantized along $\hat{u}_a(\theta'_a, \phi'_a)$. If the OX-axis lies in the scattering plane, then $\mathbf{k}_a(k_a, \theta_a, \phi_a = 0, \pi)$.

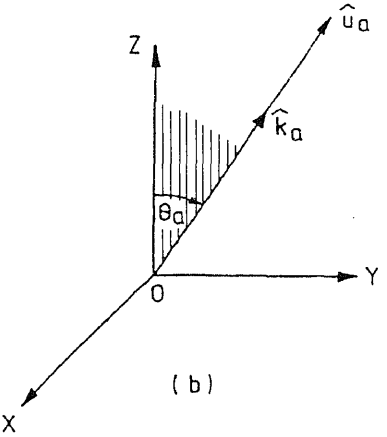


Figure 1(b). Experimental geometry for observing longitudinally polarized Auger electrons. Here $\hat{u}_a \| \hat{k}_a$ such that $\theta'_a = \theta_a$, $\phi'_a = \phi_a$.

A_{lj} to be used for atoms in eq. (II.20) are now given by

$$A_{lj}(J_f; J) = (-i)^l e^{i\sigma_l} \sqrt{(2l+1)(2j+1)} \langle J_f j | V_c | J \rangle \tag{3a}$$

with

$$A = \sum_{lj} |\langle J_f j | V_c | J \rangle|^2 \tag{3b}$$

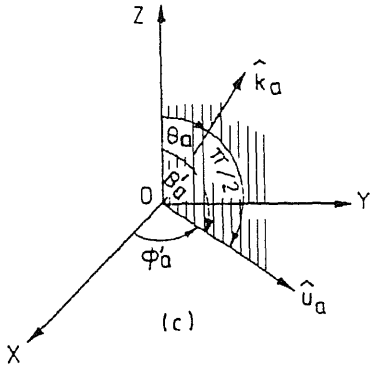


Figure 1(c). Configuration of an experiment for observing Auger electrons, with transverse polarization in the scattering plane. In this case $\theta'_a = \theta_a + \pi/2$ and $\phi' = \phi_a$ for $0 \leq \theta_a \leq \pi/2$; $\theta'_a = \theta_a - 3\pi/2$ and $\phi'_a = \phi_a + \pi$ for $\pi/2 \leq \theta_a \leq \pi$.

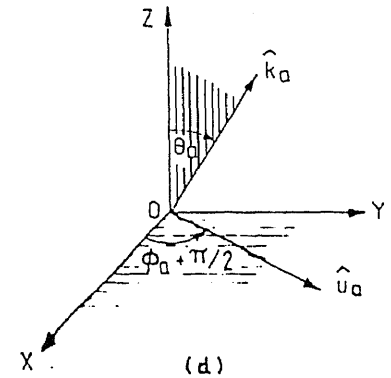


Figure 1(d). Experimental arrangement for observing Auger electrons with transverse polarization perpendicular to the scattering plane. In this geometry, $\theta'_a = \pi/2$, $\phi'_a = \phi_a + \pi/2$.

and

$$\sigma_a = K_a A \sqrt{2J+1} \langle T(J; m_r)^+_{\infty} \rangle \tag{3c}$$

is the spin-unresolved, integrated Auger current with constant K_a defined in eq. (II.2). In (3a) σ_l is the Coulomb phase [8] for the lth partial wave of Auger electron; both in (3b) and (3c), V_c is the inter-electronic Coulomb interaction. The state multipoles [9] $\langle T(J; m_r)^+_{KQ} \rangle$, present in the definition (II.13) of the normalized statistical tensor $\rho_{KQ}(J; m_r)$ needed for the angle- and spin-resolved Auger electron spectroscopy (ASRAES) expression (II.20), can be readily obtained from the expressions developed in the paper II by using the procedure explained in ref. [18]. The parameter m_r in (3c) specifies the state of polarization of the absorbed photon ($m_r = 0$ for plane and $m_r = \pm 1$ for circularly polarized light).

In order to obtain an expression for CD in atomic AES, we first need to eliminate the effects of SOI from the second step of each of the processes (1) and (2). That is, from

$(X^*, X^+$, plus Auger continuum) in (1b) and $(X^{+*}, X^{++}$, plus Auger continuum) in (2b). It is readily done by transforming the Auger decay matrix element (3a), occurring in the distribution (II.20), from the j–j to L–S coupling with the help of the relations (4.20) and (4.21) given in ref. [10]. This yields

$$\langle J_f j|V_c|J\rangle \equiv \langle (L_f S_f) J_f (l\,\tfrac{1}{2})\, j|V_c|(LS)J\rangle$$
$$= \sqrt{(2j+1)(2J_f+1)(2L+1)(2S+1)} \begin{Bmatrix} L_f & l & L \\ S_f & \frac{1}{2} & S \\ J_f & j & J \end{Bmatrix} \langle (L_f l)L|V_c|L\rangle. \tag{4}$$

Next we substitute (3a), (II.13), (3c), (4) and the geometrical factor (II.19) in eq. (II.20). The sum over j, j', J_f can then be analytically performed using an identity from ref. [11]. We then find, after some additional simplifications, that the angular distribution of spin-resolved Auger electrons emitted from atoms in the absence of SOI in the decay processes (1b) and (2b) is described by

$$\frac{\mathrm{d}\sigma_a(m_r; \hat{u}_a \mu_a)}{\mathrm{d}\hat{k}_a} = \frac{\sigma_a}{A} \sum_{\substack{ll' \\ KQ}} G(LSJ; L_f S_f; KQ; ll'; \hat{k}_a \hat{u}_a \mu_a) \rho_{KQ}(J; m_r) A_l(L_f, L)$$
$$A^*_{l'}(L_f, L). \tag{5}$$

Here the total Auger intensity σ_a is still obtained from (3c) with

$$A = \sum_l |\langle (L_f l)L|V_c|L\rangle|^2 \tag{6}$$

in place of that given by (3b). The Auger decay matrix element in L-S coupling is

$$A_l(L_f; L) = (-i)^l \mathrm{e}^{i\sigma_l} \sqrt{2l+1}\, \langle (L_f l)L|V_c|L\rangle. \tag{7}$$

Further in (5),

$$G(LSJ; L_f S_f; KQ; ll'; \hat{k}_a \hat{u}_a \mu_a) = (-1)^{1+\mu_a+L+S+L_f+S_f+K+Q}$$
$$\times (2L+1)(2S+1)\sqrt{2K+1} \sum_{\substack{L_a M_{L_a} \\ S_a M_{S_a}}} \sqrt{(2L_a+1)(2S_a+1)} \begin{pmatrix} l & l' & L_a \\ 0 & 0 & 0 \end{pmatrix}$$
$$\times \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & S_a \\ -\mu_a & \mu_a & 0 \end{pmatrix} \begin{pmatrix} L_a & S_a & K \\ M_{L_a} & M_{S_a} & -Q \end{pmatrix} \begin{Bmatrix} l & l' & L_a \\ L & L & L_f \end{Bmatrix} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & S_a \\ S & S & S_f \end{Bmatrix}$$
$$\times \begin{Bmatrix} L & S & J \\ L & S & J \\ L_a & S_a & K \end{Bmatrix} Y^{M_{L_a}}_{L_a}(\hat{k}_a) Y^{M_{S_a}}_{S_a}(\hat{u}_a) \tag{8}$$

is the geometrical factor.

The form of the distribution (5) is identical to that given in eq. (I.4) of [1] for rotating linear molecules without including either SOI or the spin-rotation interaction (SRI).

Although, the dynamical terms are naturally different in the two cases, the geometrical factors differ merely in the replacement of the respective molecular quantum numbers N and N_f in ref. [1] by the atomic total orbital angular momenta L and L_f in the present case. This means that the analysis given in the paper I for angle and spin-resolved AES of rotating linear molecules in the absence of SOI and SRI becomes immediately applicable in a similar situation (i.e., in the absence of SOI) to atoms as well.

For example, the distribution (5) for atomic targets too is completely characterized in the form of eq. (I.7) by four independent parameters. These parameters are the spin-unresolved, total Auger intensity σ_a given by eqs (3c), (6), plus β_a, γ_a, δ_a obtained from the respective equations (I.8a–c) after making the changes already mentioned in the above paragraph.

The applicability of the ASRAES (I.7) in the present case means that even CD in Auger electron spectroscopy (AES) (i.e., CDAES) of atoms can be completely described by eqs (I.13)–(I.15). This, in other words, means that CD will exist in the AES of both atomic as well as of rotating linear molecular targets in identical, both kinematical as well as physical conditions and can be treated in these two entirely different targets on an equal footing. Its important properties are:

(a) It is present in free and achiral atomic targets.
(b) It is observable only in spin-resolved integrated and differential Auger currents.
(c) No spin-dependent interations in the bound and continuum electrons are involved in either of the Auger decay processes (1b) and (2b).
(d) It does not necessarily vanish even in achiral experimental geometries [2] containing either three vectors (k_a, $\hat{u}_a$, and the direction of incidence of the CP radiation) in a single plane in eq. (I.13) or only two vectors (photon and Auger electron spin quantization axis) in eq. (I.14).
(e) Properties (I.16a–d) of CD in AES are valid for atomic targets as well.
(f) We know from eq. (II.A13) that

$$\rho_{10}(J; -m_r) = -\rho_{10}(J; m_r)$$
$$\rho_{20}(J; -m_r) = \rho_{20}(J; m_r)$$

for orientation and alignment parameters respectively. Using these, we find from the distribution (I.7)

$$\frac{d\sigma_a(-m_r; \hat{u}_a, \pm\mu_a)}{d\hat{k}_a} = \frac{d\sigma_a(m_r; \hat{u}_a, \mp\mu_a)}{d\hat{k}_a}. \tag{9}$$

The spin-resolved, integrated Auger currents (I.9a), too has properties similar to (9).
(g) CD in AES (I.13) and (I.14) is such that

$$\frac{d\sigma_a^{CD}(\hat{u}_a, -\mu_a)}{d\hat{k}_a} = -\frac{d\sigma_a^{CD}(\hat{u}_a, \mu_a)}{d\hat{k}_a} \tag{10a}$$

and

$$\sigma_a^{CD}(\hat{u}_a, -\mu_a) = -\sigma_a^{CD}(\hat{u}_a, \mu_a) \tag{10b}$$

respectively.

3. Applications

3.1 *Spin-polarization and CD in the Auger decay of photoexcited atoms*

We have calculated in paper I, the spin-resolved, integrated Auger current, its degree of polarization, and the corresponding CD in AES in terms of the total Auger current emitted in the non-radiative decay of a photoexcited rotating linear molecule. These are given in table I in ref. [1] for 102 different transitions of the type (I.1a). The results of that table are directly applicable also to the Auger electrons emitted from an atomic target through the process (1) after replacing [as explained after eq. (8)] the molecular quantum number N by the atomic total orbital angular momentum quantum number L. That is, the successive values of N, S, J given in the respective 2nd, 3rd and 4th columns of table 1 in ref. [1] will now specify the quantum numbers L, S, J of the states of photoexcited atom X^* in eq. (1).

For example, the Auger emission process (1) corresponding to the state number 7 is

$$h\nu_r + X(J_0 = \tfrac{5}{2}) \rightarrow X^*(^2P_{3/2}) \rightarrow X^+(S_f = 0) + e_a,$$

to the state number 13 is

$$h\nu_r + X(J_0 = \tfrac{1}{2}) \rightarrow X^*(^2D_{3/2}) \rightarrow X^+(S_f = 0) + e_a,$$

and to the state number 31 is

$$h\nu_r + X(J_0 = 5) \rightarrow X^*(^3F_4) \rightarrow X^+(S_f = \tfrac{1}{2}) + e_a.$$

These processes have their values of J_0 specified by the columns 8th, 6th and 8th of table I [1], respectively. We further have from table I in ref. [1] that $\gamma_a\rho_{10}(J; m_r) = m_r/8$ for each of these transitions. Here [1], $m_r = 0$, $+1$, and -1 for LP, RCP and LCP radiation of frequency ν_r absorbed in the three processes, respectively. Substitution of this value of $\gamma_a\rho_{10}(J; m_r)$ in eq. (I.9a) gives the spin-resolved, integrated Auger current $\sigma_a(m_r; \hat{u}_a, \mu_a) = (1 - (1/4)m_r\mu_a \cos\theta'_a)\sigma_a/2$; in eq. (I.9b) gives the degree $p(m_r; \hat{u}_a) = -(1/4)m_r \cos\theta'_a$ of spin-polarization of this current; in eq. (I.14) gives its circular dichroism $\sigma_a^{CD}(\hat{u}_a, \mu_a) = (1/4)\sigma_a\mu_a \cos\theta'_a$. Here σ_a is the spin-unresolved, integrated Auger current defined by (3c), (6). These values of $\sigma_a(m_r; \hat{u}_a, \mu_a)$, $p(m_r; \hat{u}_a)$, and of $\sigma_a^{CD}(\hat{u}_a, \mu_a)$ are obtained when SOI is not taken into account in the second step of any of the three above-mentioned process.

3.2 *Spin-polarization and CD in AES of Ba*

Recently, Kuntze *et al* [12–14] have studied spin-polarization of photoelectrons

$$h\nu_r + \mathrm{Ba}(5p^6 6s^2\ ^1S_0) \rightarrow \mathrm{Ba}^{+*}(5p^5 6s^2\ ^2P_{1/2,3/2}) + e_p(k_p s_{1/2}, k_p d_{3/2,5/2}) \quad (11a)$$

and of Auger electrons

$$\mathrm{Ba}^{+*}(5p^5 6s^2\ ^2P_{1/2,3/2}) \rightarrow \mathrm{Ba}^{++}(5p^6 6s^2\ ^1S_0) + e_p(k_p, p_{1/2,3/2}) \quad (11b)$$

emitted sequentially in the interaction of a barium atom with CP radiation. This obviously corresponds to the general process (2) considered in this paper with $(L_0, S_0, J_0 = 0)$,

$(L = 1; S = 1/2; J = 1/2, 3/2)$ and $(L_f, S_f, J_f = 0)$ for Ba, Ba^{+*} and Ba^{++} respectively. In order to study spin polarization of these Auger electrons and their dichroic behaviour according to the prescription explained in this paper, we take SOI into account neither in the continuum nor in the bound electrons in the second step (11b) representing the Auger decay of Ba^{+*}. It is obvious from the expression (I.8b) that the parameter γ_a is always independent of Auger dynamics. However, if $L_f = 0$, as in the present case of Ba^{++}, then the parameters β_a and δ_a too cease to depend upon the Auger decay matrix elements (see, eqs (I.18)). Thus, for such states of the excited photoion Ba^{+*} in (11b), the last three of the four parameters $(\sigma_a, \beta_a, \gamma_a, \delta_a)$ needed to describe angle and spin-resolved AES of Ba atom by eq. (I.7) become independent of the dynamical factors. The values of these three parameters are then readily calculated from eqs (I.8) (after replacing the respective molecular quantum numbers N and N_f by atomic orbital angular momenta L and L_f) to be

$$\beta_a = 0, \quad \gamma_a = 1/3\sqrt{2}, \quad \delta_a = 2\sqrt{2}/3 \tag{12a}$$

for $J = 1/2$ and

$$\beta_a = -\frac{1}{2}, \quad \gamma_a = -\frac{\sqrt{5}}{6} \quad \delta_a = -\frac{1}{3\sqrt{5}} \tag{12b}$$

for $J = 3/2$.

It is well known (see, for example, eqs (I.A8) and (II.A13)) that a state with $J = 1/2$ of a photoexcited atom or molecule, or of excited photoion of these species, will only be oriented not aligned. Kuntze *et al* [12] have given normalized orientation and alignment parameters for $Ba^{+*}(5p^5 6s^2\ ^2P_{1/2,3/2})$. The values of these parameters appropriate for present application are obtained after multiplying those given by Kuntze *et al* [12] by $\sqrt{2J+1}$. This yields

$$\rho_{10}(J = \tfrac{1}{2};\ m_r = +1) = -0.38; \quad \rho_{20}(J = \tfrac{1}{2};\ m_r = \pm 1) = 0 \tag{13a}$$

for $J = 1/2$ and

$$\rho_{10}(J = \tfrac{3}{2};\ m_r = +1) = 0.46; \quad \rho_{20}(J = \tfrac{3}{2};\ m_r = \pm 1) = 0.14 \tag{13b}$$

for $J = 3/2$. These correspond, respectively, to the photons of energies 25.8 eV and 23.5 eV incident in the process (11a) [12].

Following [12–14], let us take the reaction plane, defined in figure 1 to contain the polar axis and the propagation direction $\hat{k}_a(\theta_a, \phi_a)$ of the Auger electron, to be the X-Z plane of our photon-frame of reference. Then, obviously, the Auger electron moves out in direction $\hat{k}_a(\theta_a, \phi_a = 0)$. The three components

$$p_x = \frac{3}{4} \frac{\delta_a \rho_{10}(J; m_r) \sin 2\theta_a}{1 + \beta_a \rho_{20}(J; m_r) P_2(\cos\theta_a)}, \tag{14a}$$

$$p_y = 0, \tag{14b}$$

$$p_z = -\frac{\gamma_a - \delta_a P_2(\cos\theta_a)}{1 + \beta_a \rho_{20}(J; m_r) P_2(\cos\theta_a)} \rho_{10}(J; m_r) \tag{14c}$$

of the Auger electron polarization **p** specific to this frame of reference are readily obtained from eqs (I.10) and (I.11). Equation (14) immediately means that, if SOI is not

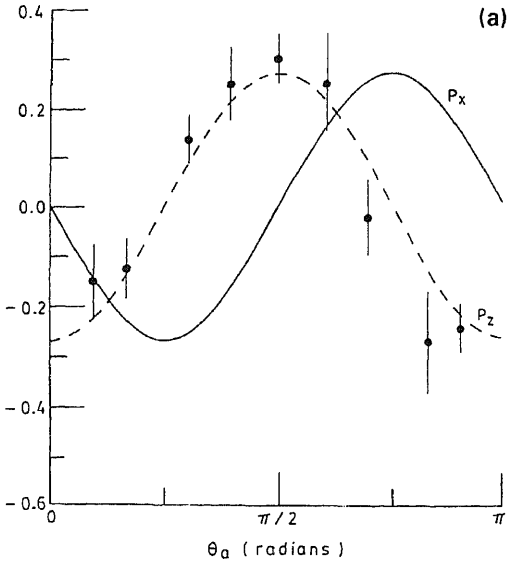


Figure 2(a). Variation of p_x and p_z components of the polarization vector of Auger electrons with the angle θ_a between directions of their propagation and of incidence of CP ionization radiation. These electrons are ejected in the non-radiative decay (11b) of excited photoion $Ba^{+*}(5p^5\,{}^2P_{1/2})$. Values shown have been calculated by substituting eqs (12a), (13a) in (14a, c). Circles (with error bars) are the values of p_z measured by Kuntze *et al* [12–14].

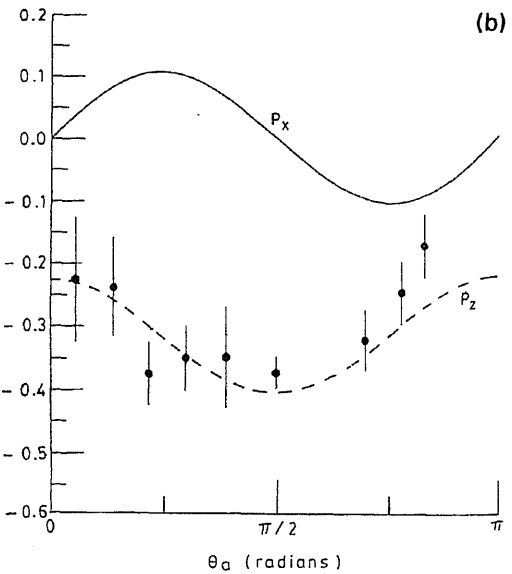


Figure 2(b). Same as figure 2(a), but for electrons ejected in the non-radiative decay (11b) of excited photoion $Ba^{+*}(5p^5\,{}^2P_{3/2})$. These values have been calculated by substituting eqs (12b), (13b) in (14a, c).

included in (11b), the Auger electrons do not have any polarization perpendicular to the scattering plane (see figure 1(d)). The last result, which was initially derived by Klar [3] in the context of atoms, is applicable to all states of photoexcited atom X^* and molecule AB^* in (1) and (I.1a) respectively, as well as of the excited atomic photoion X^{+*} in (2) and molecular photoion AB^{+*} in (I.1b,c).

Figure 2(a) contains X- and Z-components of the polarization vector **p** of Auger electrons emitted making an angle θ_a with the CP photon beam incident in the process (11b) with $J = 1/2$. These have been obtained by substituting (12a) and (13a) in (14). As both β_a and ρ_{20} vanish for $J = 1/2$, the denominator in eq. (14) is, therefore, equal to one for this case. We find from figure 2(a) that, whereas the maximum magnitude of each of the two components p_x and p_z is about 0.27, these are respectively antisymmetric and symmetric about $\theta_a = \pi/2$.

The values of the components p_x and p_z of the polarization vector **p** obtained by substituting (12b) and (13b) in (14) are shown in figure 2(b). These obviously correspond to $J = 3/2$ in the Auger transition (11b). Here too we find that, similar to figure 2(a), p_x and p_z are respectively antisymmetric and symmetric with respect to $\theta_a = \pi/2$. Moreover, $|p_x(J = 1/2)|_{max} > |p_x(J = 3/2)|_{max}$, but $|p_z(J = 1/2)|_{max} < |p_z(J = 3/2)|_{max}$. The most significant effects on polarization of Auger electrons of an increase in the total angular momentum of Ba^{+*}, which results both in orientation as well as in alignment of $p_{3/2}$ vacancy, are that (i) the maximum magnitude of $p_x(J = 3/2)$ is much smaller than that of $p_x(J = 1/2)$, $p_z(J = 1/2)$, and of $p_z(J = 3/2)$; (ii) $p_z(J = 3/2) < 0$ for all directions of propagation of Auger electrons in the X-Z plane (i.e., for all values of θ_a).

In order to see the effect of SOI in the second step of the process (11), we have calculated the parameters β_a, γ_a, δ_a, and ξ_a also from eqs (II.22). As $J_f = 0$ and a single p-partial wave of Auger electron contributes in the process (11b), all of the four parameters in eqs (II.22) become independent of Auger dynamics.

We find in this case that, while $\xi_a = 0$, each of the β_a, γ_a, and δ_a have values equal to those given by (12a) and (12b) for $J = 1/2$ and $3/2$, respectively. Consequently, with normalized orientation and alignment parameters taken from (13a) for $J = 1/2$ and from (13b) for $J = 3/2$ states of Ba^{+*}, the Cartesian components of the polarization vector of Auger electron obtained from eq. (II.27) also become identical to those given in (14) in this paper and shown in figure 2(a) and 2(b) herein.

Kuntze *et al* have measured p_y [13, 14] and p_z [12–14] both for $J = 1/2$ and $3/2$ in the process (11). These measurements [12–14] of p_z have been shown by open circles (along with error bars) in figure 2 in the present paper. On comparing the measured values [12–14] of p_z with those calculated in the present paper (and shown by the continuous curve in figure 2), we find that there is good agreement between the two, both in shape as well as magnitude. The same is true for p_y which has been measured [13, 14] to be zero. This agreement between theory and experiment shows that if the dynamics does not contribute to a particular Auger decay process, then the theoretical results obtained without SOI agree with those calculated and measured taking SOI into account.

Figures 3(a) and 3(b) contain the $\sigma_a^{CD}(\hat{u}_a, \mu_a)/\sigma_a$ in the spin-resolved, integrated Auger current produced in the respective decay of $J = 1/2$ and $J = 3/2$ states of Ba^{+*} in the process (11b). These have, respectively, been calculated by substituting (12) and (13) in (I.14). In each of the figures 3, curve 1 corresponds to the projection $-(1/2)\hbar$ and curve 2 to $+(1/2)\hbar$ of the spin-angular momentum of the Auger electron along the quantization

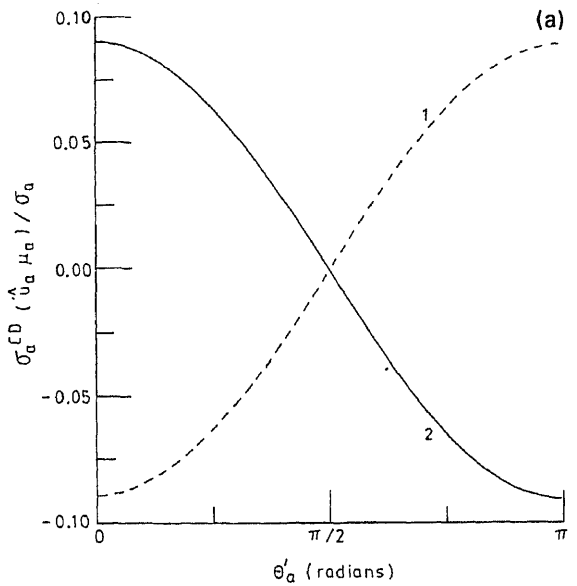


Figure 3(a). Variation of the ratios of CD in spin-resolved, integrated Auger current to the total (spin-unresovled) Auger intensity with the polar angle θ'_a between the spin-quantization direction of Auger electrons and the CP photon beam. These electrons are ejected in the non-radiative decay (11b) of the excited photoion $Ba^{+*}(5p^5\,{}^2P_{1/2})$. Ratios in this figure are calculated by substituting eqs (12a), (13a) in (I.14). Curve 1: $\mu_a = -1/2$; Curve 2: $\mu_a = +1/2$.

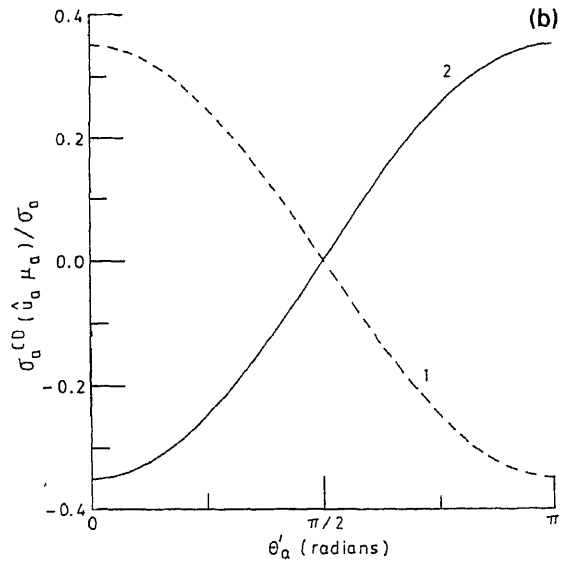


Figure 3(b). Same as figure 3(a), but for electrons emitted in the non-radiative decay (11b) of the excited photoion $Ba^{+*}(5p^5\,{}^2P_{3/2})$. Ratios in this figure are calculated by substituting eqs (12b), (13b) in (I.14).

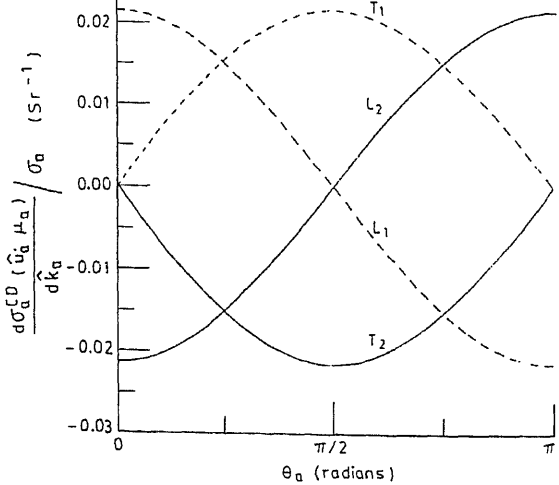


Figure 4(a). Variation of the ratios of CD in spin-resolved, differential Auger current to the total (spin-unresolved) Auger intensity with the angle θ_a between the outgoing electron and the CP radiation. These electrons are emitted in the non-radiative decay (11b) of $Ba^{+*}(5p^5\,{}^2P_{1/2})$. Curves $L_1(\mu_a = -1/2)$ and $L_2(\mu_a = +1/2)$ are calculated by substituting eqs (12a), (13a) in (I.16a) for longitudinal polarization of Auger electrons. Curves $T_1(\mu_a = -1/2)$ and $T_2(\mu_a = +1/2)$ are obtained by substituting eqs (12a), (13a) in (I.16b) for transverse polarization of Auger electrons in the reaction plane. (Here $\mu_a = \pm 1/2$ is the projection of Auger electrons spin along the quantization direction $\hat{u}_a$.)

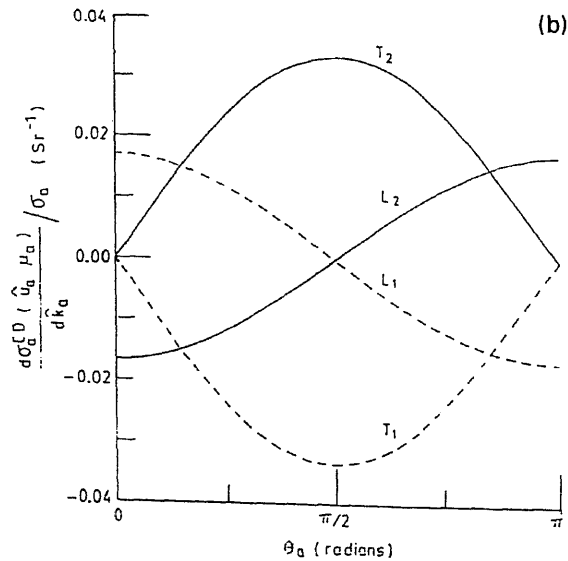


Figure 4(b). Same as figure 4(a) but for Auger electrons emitted in the non-radiative decay (11b) of the excited photoion $Ba^{+*}(5p^5\,{}^2P_{3/2})$. These ratios are calculated by substituting (12b), (13b) in (I.16a) for L_1 and L_2 and in (I.16b) for T_1, T_2.

ch figure obviously obey the property (10b). Although, the magnitude of $\sigma_a^{CD}(\hat{u}_a, \mu_a)/\sigma_a$ each of the figures 3 is within the experimentally measurable limit, the CD for $J = 3/2$ figure 3(b)) is almost four times that when $J = 1/2$ (in figure 3(a)) state of Ba^{+*}.

The ratio of the CD in spin-resolved differential Auger current (i.e., $d\sigma_a^{CD}(\hat{u}_a, \mu_a)/d\hat{k}_a$) that of the total Auger intensity (σ_a) for different experimental geometries is shown in ure 4. Figure 4(a) corresponds to the decay of $Ba^{+*}(^2P_{1/2})$ state and is obtained by stituting (12a), (13a) in (I.16); whereas, figure 4(b) is for the Auger electrons emitted m $Ba^{+*}(^2P_{3/2})$ calculated by combining (12b), (13b) and (I.16). Curves $\mu_a = -1/2)$ and $L_2(\mu_a = +1/2)$ in both of these figures are for longitudinally larized Auger electrons corresponding to eq. (I.16a). Similarly, the remaining two rves $T_1(\mu_a = -1/2)$ and $T_2(\mu_a = +1/2)$ in each of the figures 4(a) and 4(b) give us in spin-resolved differential Auger current for transverse experimental arrangement eq. (I.16b). That is, L_1 and L_2 are for $\hat{u}_a \| \hat{k}_a$; whereas, T_1 and T_2 are for $\hat{u}_a \perp \hat{k}_a$. These cases correspond to the experimental geometries shown in figures 1(b) and 1(c), pectively, wherein the Auger electron spin quantization direction $\hat{u}_a$ and its pagation vector $\hat{k}_a$ are in the plane which also contains the CP incident photon beam. rthermore, all the curves in the figures 4(a) and 4(b) satisfy the property (10a). This ans that, for all values of the angle θ_a between the direction of ejection of Auger ctron and of incidence of CP radiation, L_1 and T_1 are always negative of L_2 and T_2 pectively. Finally, the ratios of the CD in spin-resolved differential Auger current to al Auger intensity in the figures 4(a) and 4(b) are well within the measurable limits of sent experimental facilities.

Spin-polarization and CD in AES of Xe

e following photoionization process

$$h\nu_r + \mathrm{Xe}(4d^{10}5s^25p^6\,{}^1S_0) \rightarrow \mathrm{Xe}^{+*}(4d^95s^25p^6\,{}^2D_{5/2}) + e_p(k_p\, p_{1/2}, k_p\, f_{5/2,7/2}) \tag{15a}$$

Xe succeeded by the spontaneous, non-radiative decay

$$\mathrm{Xe}^{+*}(4d^95s^2\,5p^6\,{}^2D_{5/2}) \rightarrow \mathrm{Xe}^{++}(4d^{10}5s^05p^6\,{}^1S_0) + e_p(k_a d_{5/2}) \tag{15b}$$

the excited photoion Xe^{+*} has been studied by several workers, both theoretically as ll as experimentally. More recently, Kammerling and Schmidt [15] have measured e-photon two-step, double ionization for (15) and theoretically studied [15, 16] both gular and spin-correlation between photo- and Auger-electrons emitted sequentially in above process.

The values of the three polarization parameters obtained from eqs (I.8) for this case is

$$\beta_a = -\frac{2}{\sqrt{21}}, \quad \gamma_a = -\frac{1}{3}\sqrt{\frac{7}{10}}, \quad \delta_a = -\frac{2}{3}\sqrt{\frac{2}{35}}. \tag{16a}$$

ese parameters are independent of the Auger decay dynamics because we have, in the sent case $L_f = 0$ and the Auger electron is represented by a single $d(l = 2)$-partial

wave in eq. (15b). Since $J_f = 0$ as well, eqs (II.22) which includes SOI in the Auger decay process (15b) will also yield the values of β_a, γ_a, δ_a given in (16a), with $\xi_a = 0$ (obtained from eq. (II.22d)). Thus similar to the case of Ba, discussed earlier, the SOI does not contribute to the second step in double ionization (15) of Xe.

In order to calculate the components of the polarization vector of Auger electrons and CD for (15b), we need to know both orientation and alignment parameters $\rho_{10}(J; m_r)$ and $\rho_{20}(J; m_r)$, respectively. These have been calculated herein using the photoionization matrix elements given by Kammerling and Schmidt [15]. The expression used in ref. [15] for the integrated cross-section for photoionization (15a) is [16]

$$\sigma(4d_{5/2}) = \frac{8\pi^4}{\omega c}\,(|d'(p_{3/2})|^2 + |d'(f_{5/2})|^2 + |d'(f_{7/2})|^2).$$

Here ω is the energy of the photon incident in (15a) and c (= 137) is the speed of light in atomic unit (a.u.); d' are the photoionization amplitudes in electric dipole ($E1$) approximation. But the theoretical frame-work developed in ref. [1] and [17] gives the integrated photoionization cross-section for (15a) to be

$$\sigma(4d_{5/2}) = K_p(|d(p_{3/2})|^2 + |d(f_{5/2})|^2 + |d(f_{7/2})|^2),$$

with $K_p = 3\pi(e^2/\alpha_0 E_r^2)$ in absolute units. Here e is the electronic charge, α_0 the dimensionless fine structure constant and E_r the photon energy. Converting it to atomic units, we get $K_p = 3\pi c^2/\omega^2$. Therefore, in order to be able to use the $E1$ amplitudes given by Kammerling and Schmidt [15], we should first multiply them by $8\pi^3\omega/(3c^3)$. This 're-normalization' procedure will, however, only change the magnitude of $E1$ photoionization matrix elements, but their [15] phases will remain unaltered. Therefore, the magnitude of the $E1$ transition amplitudes we have used in calculating the orientation and alignment parameters for photoionization step (15a) are

$$d(p_{3/2}) = 1.459 \times 10^{-3}\ \text{a.u.},$$
$$d(f_{5/2}) = -1.385 \times 10^{-3}\ \text{a.u.},$$
$$d(f_{7/2}) = -5.010 \times 10^{-3}\ \text{a.u.},$$

with associated phases given in [15]. The corresponding normalized orientation and alignment factors for $Xe^{+*}(4p^9\,{}^2D_{5/2})$ are then readily obtained from (II.13) to be [18]:

$$\rho_{10}(J;\ m_r) = -1.86 m_r; \quad \rho_{20}(J;\ m_r) = 0.279(3m_r^2 - 2). \tag{16b}$$

Here [1], [17], $m_r = 0$ for LP and $m_r = \pm 1$ for CP light incident in (15a). The energy of the LP photon in experiment of [15] was 94.5 eV.

The X- and Z-components of the polarization vector **p** of Auger electrons emitted from Xe in the process (15) are shown in figure 5. The results have been computed by substituting (16) in (14). On comparing figures 2(b) and 5 we find that the nature of variation of p_x and p_z with the angle of emission θ_a of Auger electrons in the two figures is almost the same. It is because both orientation and alignment contribute to the polarization of Auger electrons in the non-radiative decay of $Ba^{+*}(^2P_{3/2})$ in (11b) and of $Xe^{+*}(^2D_{5/2})$ in (15b). However, the magnitude of p_z for Xe-Auger electron is much larger than for those coming out from Ba^{+*}, but p_x has almost equal magnitude in

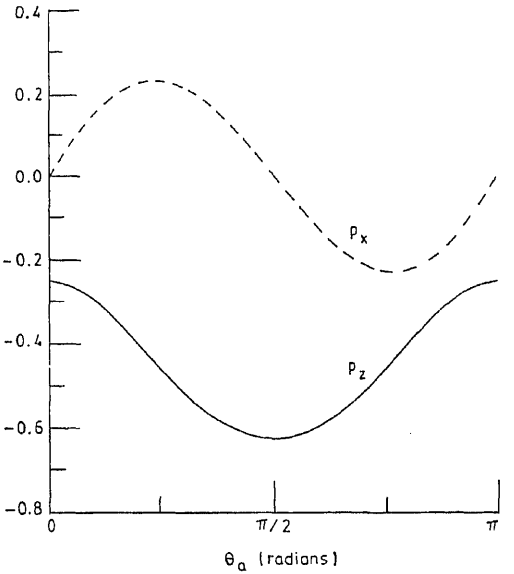


Figure 5. Same as figure 2(a) but for Auger electrons emitted in the non-radiative decay (15b) of excited photoion $Xe^{+*}(4p^9\,{}^2D_{5/2})$. These values have been calculated by substituting eq. (16) in (14a, c).

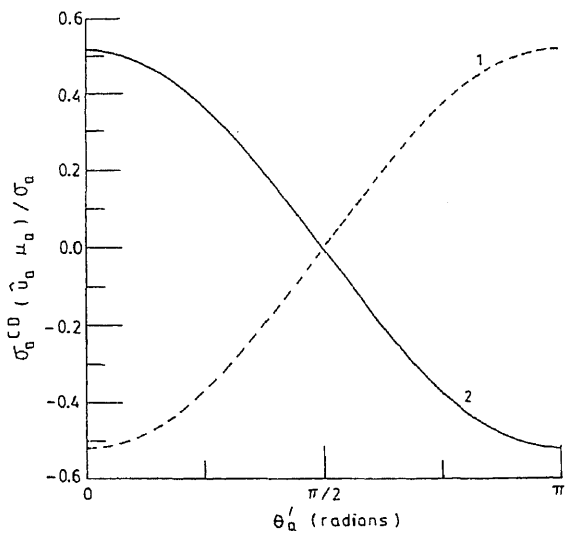


Figure 6. Same as figure 3(a) but for Auger electrons emitted in the non-radiative decay (15b) of excited photoion $Xe^{+*}(4p^9\,{}^2D_{5/2})$. These ratios are calculated by substituting eqs (16) in (I.14).

figures 2(a) and 5, which is almost twice of that given in figure 2(b) for $Ba^{+*}({}^2P_{3/2})$. This simply means that the degree of spin-polarization of Auger electrons emitted in (15) is expected to be much larger than those coming out from (11).

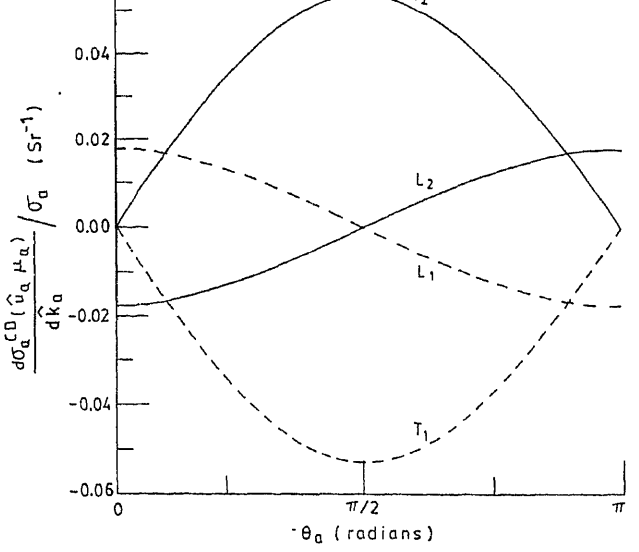


Figure 7. Same as figure 4(a) but for Auger electrons emitted in the non-radiative decay (15b) of excited photoion $Xe^{+*}(4p^9\,{}^2D_{5/2})$. These ratios are calculated by substituting eqs (16) in (I.16a) for L_1 and L_2 and in (I.16b) for T_1, T_2.

As far as we know, unfortunately, no experimental measurements or any other calculations for spin-polarization of Auger electrons emitted in the process (15) in Xe are available. Talkki *et al* [19] have computed those parameters which are needed in eq. (II.27b) to calculate p_y component of the polarization vector perpendicular to the reaction plane in the presence of SOI. But, for the transition (II.27b) in Xe^{+*}, p_y has already been shown to vanish identically elsewhere in this paper. Consequently, the results presented in figure 5 are the first theoretical prediction of the components p_x and p_z of the polarization vector **p** of Auger electrons emitted in the process (15) in Xe.

In the respective figures 6 and 7, we have shown the ratios of the CD in spin-resolved integrated and differential Auger currents to the total Auger intensity for the process (15). These have been calculated by substituting eq. (16) in (I.14) and (I.16), respectively. In figure 7, both longitudinal (L_1, L_2) as well as transverse (T_1, T_2) geometries, corresponding to eqs (I.16a) and (I.16b) respectively, have been considered. (These experimental arrangements are shown in the respective figures 1(b) and 1(c).) The magnitude of the ratios in these figures are usually larger than those of Ba shown in figures 3 and 4. Another interesting comparison between CD in Ba and Xe AES is that variations in figure 6 with respect to θ_a' and in figure 7 with respect to θ_a for $Xe^{+*}(^2D_{5/2})$ are similar to those found for $Ba^{+*}(^2P_{3/2})$ in figures 3(b) and 4(b), respectively.

4. Conclusion

In this paper, we have shown that the same set of equations describe angle- and spin-resolved AES following photoabsorption in both atoms as well as in rotating linear

molecules with or without including SOI in the second step of Auger emission. Although, the dynamical terms needed for atomic systems are naturally different from those for molecules, the geometrical factor, is however, identical in the two cases. We have, further, shown that CD exists in the AES of both of these targets in almost identical physical as well as geometrical conditions and is, again, represented by the same equations which do not include SOI in the Auger decay dynamics. We have used these equations for calculating components of the polarization vector and CD for Auger electrons emitted from the excited photoions $Ba^{+*}(5p^5\,{}^2P_{1/2,3/2})$ and $Xe^{+*}(4p^9\,{}^2D_{5/2})$.

Our investigations show that if the total orbital angular momentum L_f as well as the total angular momentum J_f of the doubly charged ion left after Auger emission is zero (i.e., it is in its $^1S_0$ state) and the departing electron is represented by a single partial wave, then the degree of spin polarization of Auger electrons is not affected by the absence or presence of SOI in bound and/or continuum electrons. The Auger electrons in this case are spin-resolved if and only if their emission is preceeded by the absorption of CP light, whether or not SOI is included. In such experiments, polarization of Auger electrons is confined only to the plane which contains both the outgoing electrons as well as the CP ionizing radiation, and vanishes identically in a direction perpendicular to this plane. In the three examples considered in this paper, the maximum magnitude of the p_z component of polarization vector along the photon beam is either equal to or larger than its p_x component perpendicular to the photon beam. Moreover, the degree of spin polarization of Auger electrons produced in the decay of $Xe^{+*}(4p^9\,{}^2D_{5/2})$ is much larger than of those emitted from either of the two states $(5p^5\,{}^2P_{1/2,3/2})$ of Ba^{+*}. The CD investigations performed here are useful for studying spin-resolved Auger spectroscopy, and for investigating the influence of SOI on Auger electrons. The circular dichroism calculated in this paper in the AES of Ba and Xe suggests that it is sufficiently larger and can be studied experimentally. In addition, being the difference of two aspects of almost equal values, CD studies provide more stringent test for theoretical models.

Acknowledgements

The author is thankful to the Council of Scientific and Industrial Research for the award of a Research Associateship.

References

[1] N Chandra and S Sen, *Phys. Rev.* **A48**, 2084 (1993) [This paper will henceforth be referred to as I and an equation number of this paper cited herein is enclosed in the parentheses with the letter I]
[2] G Schonhense, *Phys. Scr.* **T31**, 255 (1990)
[3] H Klar, *J. Phys.* **B13**, 4741 (1980)
[4] N M Kabachnik, *J. Phys.* **B14**, L377 (1981)
[5] K N Huang, *Phys. Rev.* **A26**, 2274 (1982)
[6] K Blum, B Lohmann and E Taute, *J. Phys.* **B19**, 3815 (1986)
[7] N M Kabachnik and O V Lee, *J. Phys.* **B22**, 2705 (1989)
[8] N Chandra, *Chem. Phys.* **108**, 301 (1986)
[9] K Blum, *Density matrix theory and applications* (Plenum, New York, 1981)
[10] R N Zare, *Angular momentum* (John Wiley, New York, 1988)

[11] D A Varshalovich, A N Moskalev and V K Khersonskii, *Quantum theory of angular momentum* (World Scientific, Singapore, 1988) [eq. (44) on page 474]
[12] R Kuntze, M Salzmann, N Bowering and U Heinzmann, *Phys. Rev. Lett.* **70**, 3716 (1993)
[13] R Kuntze, M Salzmann, N Bowering and U Heinzmann, *Z. Phys.* **D30**, 235 (1994)
[14] R Kuntze, M Salzmann, N Bowering, U Heinzmann, V K Ivanov and N M Kabachnik, *Phys. Rev.* **A50**, 489 (1994)
[15] B Kammerling and V Schmidt, *Phys. Rev. Lett.* **67**, 1848 (1991)
[16] B Kammerling and V Schmidt, *J. Phys.* **B26**, 1141 (1993)
[17] N Chandra and S Sen, *J. Chem. Phys.* **98**, 5242 (1993) [This paper will hence forth be referred to as II. An equation number of this paper cited herein is enclosed in the parentheses with letters II]
[18] In ref. [17], we have derived respective expression [(A11), (A11′)] and [(A13), (A13′)] for state multiples and normalized statistical tensors for photoion of a linear molecule rotating according to the Hund's coupling scheme (a) or (b). These expressions are applicable to the present atomic studies as well provided the following $\langle J;\ jl|F(j_t)|J_0;\ 1\rangle = (2J_0+1)^{-1/2}\sum_{J_T}(2J_T+1)\begin{Bmatrix} j & J & J_T \\ 1 & J_0 & j_t \end{Bmatrix}\langle J;\ jl|F(J_T)|J_0;\ 1\rangle$ definition, in place of that given in eq. (II.A10), of 'reduced' amplitude is used in eqs (II.A9)–(II.A13′). Here, the $E1$ ionization amplitude is taken from eq. (9) in N Chandra, *Phys. Rev.* **A42**, 4050 (1990)
[19] J Tulkki, N M Kabachnik and H Aksela, *Phys. Rev.* **A48**, 1277 (1993)

—journal of
physics

November 1997
pp. 479–491

Two Coulomb waves theory for direct excitation

SIMA CHAKRABARTI and DURGA P DEWANGAN
Physical Research Laboratory, Navrangpura, Ahmedabad 380 009, India
Email: dewangan@prl.ernet.in

MS received 30 January 1997; revised 6 August 1997

Abstract. A theoretical model of Dewangan, in which the total scattering wave function is approximated by a distorted wave containing two Coulomb wave functions, is discussed and its relation with the Brauner–Briggs–Klar model for ionization is examined. An important feature of the theory is that it includes a second Born amplitude naturally and in addition, contains, albeit approximately, both real and imaginary parts of all higher order Born terms. The theory is applied to study the $1s \rightarrow 2s$ excitation of hydrogen by electrons in the energy range 54.4 to 400 eV. The differential and integral cross sections predicted by the theory are compared with the results of other theories and experimental data at 54.4 eV and a good agreement is found.

Keywords. Brauner–Briggs–Klar model; two-Coulomb waves.

PACS Nos 34.80; 34.90

1. Introduction

In recent years, the Brauner–Briggs–Klar (BBK) scattering wave function [1, 2] has received considerable attention in the study of the ionization of atoms. The BBK wave function for electron/positron-hydrogen scattering essentially consists of a product of three Coulomb wave functions, and properly satisfies the boundary conditions. The origin of the three Coulomb wave functions is as follows. A product of two Coulomb wave functions is used to describe the scattered electron interacting with the two Coulomb potentials produced by the ejected electron and the target nucleus. The third Coulomb wave function is the continuum eigenfunction of the ejected electron in the field of the target nucleus. Thus, the BBK wave function describes the motion in the final channel of the ejected electron and the projectile interacting with each other and also with the target nucleus. The approximate scattering amplitude is obtained by replacing the exact scattering wave function by the BBK wave function in the prior form of the transition amplitude. An extension of the model to ionization of helium has been given by Franz and Altick [3], and Berakdar and Briggs [4]. This model has been applied recently by Berakdar *et al* [5] to study the orientation and alignment effects in ionization of H($2p$) by electron impact. The BBK wave function has also been successful in the investigation of angular distribution of electrons following double photoionization [6, 7].

In view of the current interest in the BBK theory for ionization and photoionization, it is of considerable importance to examine its applicability to direct excitation, since this would give information about its theoretical structure. In this work, we use an

approximate total scattering wave function for excitation of hydrogen proposed by Dewangan [8] and discuss its relation with the BBK wave function for ionization. We extend our preliminary study of the $1s \rightarrow 2s$ excitation of hydrogen by electron impact, performed at a single energy of 54.4 eV, to 54.4–400 eV covering a wide range in the intermediate energy region and examine the theoretical predictions by comparing them with the results of other current theories. We use atomic units, unless otherwise stated.

2. The two Coulomb wave theory for direct excitation

We consider the collision of an electron with a hydrogen atom. Let $\mathbf{r}$ and $\mathbf{r}_0$ denote the position vectors of the target electron and the incident electron with respect to the target nucleus (assumed to be infinitely heavy) respectively, and let $\mathbf{r}_{01} = \mathbf{r} - \mathbf{r}_0$ be the position vector of the incident electron with respect to the target electron. We denote the final wave vector of the incident electron by $\mathbf{k}_f$, that of the ejected electron by $\mathbf{k}$ and define $\mathbf{k}_{01} = (\mathbf{k} - \mathbf{k}_f)/2$. The BBK wave function for the final (ionization) channel is given by [1, 2]

$$\psi_f^- = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}_f\cdot\mathbf{r}_0 + i\mathbf{k}\cdot\mathbf{r}} \Gamma(1 - i\alpha_{eT}) e^{-\pi\alpha_{eT}/2} {}_1F_1(i\alpha_{eT}, 1; -ikr - i\mathbf{k}\cdot\mathbf{r})$$
$$\times \Gamma(1 - i\alpha_{PT})\, e^{-\pi\alpha_{PT}/2}\, {}_1F_1(i\alpha_{PT}, 1; -ik_f r_0 - i\mathbf{k}_f\cdot\mathbf{r}_0)$$
$$\times \Gamma(1 - i\alpha_{eP})\, e^{-\pi\alpha_{eP}/2}\, {}_1F_1(i\alpha_{eP}, 1; -ik_{01} r_{01} - i\mathbf{k}_{01}\cdot\mathbf{r}_{01}), \quad (1)$$

where $\alpha_{eT} = -1/k$, $\alpha_{PT} = Z_P/k_f$ and $\alpha_{eP} = -Z_P/|\mathbf{k} - \mathbf{k}_f|$ in which $Z_P = -1$ is the charge of the incident electron. The BBK transition amplitude for ionization is obtained by approximating the exact total scattering wave function with the above wave function in the prior form of the transition amplitude.

One can extend the BBK model to excitation of hydrogen in a straightforward manner. The BBK scattering wave function for excitation is obtained by replacing $\Gamma(1 - i\alpha_{eT})\, e^{-\pi\alpha_{eT}/2}\, {}_1F_1(i\alpha_{eT}, 1; -ikr - i\mathbf{k}\cdot\mathbf{r})$, the Coulomb eigenfunction for the ejected electron, by $\phi_f(\mathbf{r})$, the bound atomic eigenfunction of the excited final state, and putting $\mathbf{k} = 0$,

$$\psi_f^- = e^{i\mathbf{k}_f\cdot\mathbf{r}_0}\phi_f(\mathbf{r})|\Gamma(1 + i\nu_f)|^2 {}_1F_1(-i\nu_f, 1; -ik_f r_0 - i\mathbf{k}_f\cdot\mathbf{r}_0)$$
$$\times {}_1F_1(i\nu_f, 1; -ik_f r_{01}/2 + i\mathbf{k}_f\cdot\mathbf{r}_{01}/2), \quad (2)$$

where we have defined $\nu_f = Z_P/k_f$. The BBK version of the transition amplitude for excitation is obtained by using this wave function in place of the exact wave function in the prior form of the transition amplitude.

In the context of excitation, however, it is traditional to consider the post form of the transition amplitude, rather than the prior form which is used in ionization. For the reasons that will become clear in the following, we replace $k_f/2$ by k_f in the argument of the confluent hypergeometric function appearing in (2). If we denote by $\phi_i(\mathbf{r})$ the bound atomic wave function of the initial state and by $\mathbf{k}_i$ the incident wave vector, we can immediately write the post form of scattering wave function (2) corresponding to the

entrance channel as [8]

$$\psi_i^+ = \phi_i(\mathbf{r})e^{i\mathbf{k}_i\cdot\mathbf{r}_0}|\Gamma(1+i\nu)|^2\,{}_1F_1(-i\nu, 1; ik_ir_0 - i\mathbf{k}_i\cdot\mathbf{r}_0)$$
$$\times{}_1F_1(i\nu, 1; ik_ir_{01} + i\mathbf{k}_i\cdot\mathbf{r}_{01}), \tag{3}$$

where $\nu = Z_P/k_i$. The wave function (3) contains, in essence, a product of two Coulomb wave functions which describes the continuum state of the incident electron interacting simultaneously with both the target nucleus and the bound target electron. By using the asymptotic forms of the confluent hypergeometric functions [9], we may easily see that the wave function (3) satisfies proper boundary conditions

$$\psi_i^+ \sim \phi_i(\mathbf{r})e^{i\mathbf{k}_i\cdot\mathbf{r}_0} + \text{scattered wave}. \tag{4}$$

In the T-matrix formalism, when the exact scattering wave function is approximated by (3), the post form of the transition amplitude for the $\phi_i(\mathbf{r}) \rightarrow \phi_f(\mathbf{r})$ excitation is given by

$$T_{fi} = \int d\mathbf{r}\, d\mathbf{r}_0\, e^{-i\mathbf{k}_f\cdot\mathbf{r}_0}\phi_f^*(\mathbf{r})\left(\frac{Z_P}{r_0} - \frac{Z_P}{r_{01}}\right)\psi_i^+. \tag{5}$$

Since the continuum state of the incident electron is described, in essence, by two Coulomb wave functions, the present theory, for brevity, will be called the two Coulomb waves (TCW) theory.

It should be mentioned here that the use of the wave function of the form (3) in atomic scattering calculations is not new. Vainshtein, Presnyakov and Sobelman (VPS) [10] in the context of excitation of hydrogen by electron impact, and Cheshire [11], in the context of electron capture in proton-hydrogen collisions (the CDW method), were the first to propose the use of a product of two Coulomb wave functions to describe the distortion effects on the scattering wave function. In the study of electron impact excitation of hydrogen, Vainshtein *et al* [10] used the post form of the transition amplitude (the VPS approximation). But, to evaluate the amplitude, Vainshtein *et al* [10] and subsequently other authors (see the references given in [12, 13]) applied peaking approximations to simplify the calculations. Dewangan and Bransden [12] analysed the physical content of the VPS approximation for excitation of hydrogen by proton impact and showed (see below) that the application of the peaking approximation to the VPS amplitude in essence leads to a first order amplitude that is a poor approximation to the VPS amplitude. Their analysis showed that the contribution to the VPS scattering amplitude from the incident proton-electron matrix element contains a term which is identical to a second Born amplitude for this interaction computed in a closure approximation. However, the matrix element of the incident proton-target nucleus interaction that occurs in the VPS amplitude, contains a term which is not identical to, but has a structure similar to a second Born term, for this interaction. Following this work, Dewangan [8], by modifying the VPS wave function, wrote down the TCW wave function (3) as well as the TCW transition amplitude (5) for electron hydrogen scattering and demonstrated that the transition amplitude (5) contains full closure approximation second Born terms corresponding to the incident electron-target electron and the incident electron-target nucleus interaction. Dewangan and Khadkikar [14] have developed a distorted wave expansion for the transition amplitude whose first order term is the TCW amplitude (5). It is shown in that paper that the TCW amplitude (5) contains (1) a second Born term, (2)

both real and imaginary parts of all higher (≥ 3) order Born terms in an approximate way, and that (3) the well known Glauber amplitude (for a review see [15, 16]) can be derived by introducing further approximations in the TCW amplitude. Like the TCW amplitude, the Glauber amplitude contains contributions from all orders ($n \geq 2$) of the Born terms, but in contrast to the TCW model, either all its odd-order terms are purely real and all its even-order terms are purely imaginary or vice versa. Evidently, the TCW theory includes contributions from all orders of the Born terms and goes beyond the Glauber approximation.

The TCW wave function (3) is very similar in spirit to the VPS wave function (or the wave function of the continuum distorted wave (CDW) model of Cheshire [11], see Dewangan and Eichler [17] for a discussion) but differs from the VPS (or the CDW) wave function in that the latter is based on the use of the Jacobi coordinates [18]. For an explicit expression of the VPS (or the CDW) wave functions and the transition amplitudes in both the post and prior forms for electron-hydrogen scattering, the reader is referred to Sinha *et al* [13].

One can show by following the method of Dewangan and Bransden [12] that the wave function (2), unlike (3), does not reproduce the full closure approximation second Born term and therefore we use only the wave function (3). Dewangan [8] also made a preliminary application of the TCW model to the $1s \rightarrow 2s$ excitation of hydrogen by electrons at 54.4 eV and indeed, to our knowledge, except for this application, no other calculations without further approximations for direct excitation in this method have been reported. However, as the present work based on improved numerical techniques shows, due to inaccuracies in performing numerical quadratures, the differential cross sections calculated in [8] were in error by up to about 15%, depending on scattering angles. Here it must be mentioned that Chen *et al* [19] have applied the TCW theory without any further approximation to study Ps($1s$) formation in e^+–H($1s$) collisions.

In order to perform numerical computation, we follow the method of Dewangan [8] and transform the transition amplitude (5) into momentum space. Defining the form factor,

$$G_{\rm fi}(\mathbf{p}) = \int \mathrm{d}\mathbf{r}\, \mathrm{e}^{i\mathbf{p}\cdot\mathbf{r}} \phi_{\rm f}^*(\mathbf{r}) \phi_{\rm i}(\mathbf{r}), \tag{6}$$

the transition amplitude (5) may be reduced to

$$T_{\rm fi} = T_{\rm fi}^{(1)} + T_{\rm fi}^{(2)}, \tag{7}$$

where $T_{\rm fi}^{(1)}$ is given by

$$T_{\rm fi}^{(1)} = T_{\rm fi}^{\rm B1} \left(|\Gamma(1+i\nu)|^2 \, {}_2F_1 \left[-i\nu, i\nu; 1, \left(\frac{\mathbf{k}_i \cdot \mathbf{q}}{\omega_{\rm fi}} \right)^2 \right] \right). \tag{8}$$

Here, $\omega_{\rm fi} = \epsilon_{\rm f} - \epsilon_{\rm i}$, with $\epsilon_{\rm i}$ and $\epsilon_{\rm f}$ being the eigenenergies of the initial and final atomic states, respectively, $\mathbf{q} = \mathbf{k}_i - \mathbf{k}_f$ is the momentum transfer and $T_{\rm fi}^{\rm B1}$ is the usual first Born transition amplitude given by

$$T_{\rm fi}^{\rm B1} = Z_P \frac{4\pi}{q^2} [G_{\rm fi}(0) - G_{\rm fi}(\mathbf{q})]. \tag{9}$$

Clearly, the amplitude $T_{\rm fi}^{(1)}$, eq. (8), is in essence a first order approximation.

The amplitude $T_{\rm fi}^{(2)}$ may be written as

$$T_{\rm fi}^{(2)} = Z_P|\Gamma(1+i\nu)|^2\left(\frac{4\nu k_i}{\pi}\right)[T^N + T^E]. \tag{10}$$

Here, the matrix element T^N is the contribution from the interaction of unit strength $(+1/r_0)$, and is given by

$$T^N = \lim_{\alpha\to 0,\epsilon\to 0}\int \mathrm{d}\mathbf{p}\,\frac{[G_{\rm fi}(\mathbf{k}_i-\mathbf{p})-G_{\rm fi}(0)][|\mathbf{p}-\mathbf{k}_i-\mathbf{k}_f|^2+(\alpha-ik_i)^2]^{+i\nu}}{|\mathbf{p}-\mathbf{k}_i|^{(2-2i\nu)}|\mathbf{p}-\mathbf{k}_f|^{(2+2i\nu)}(p^2-k_i^2-i\epsilon)^{(1+i\nu)}}. \tag{11}$$

The term T^E is the contribution from the interaction of unit strength $(+1/r_{01})$,

$$T^E = \lim_{\alpha\to 0,\epsilon\to 0}\int \mathrm{d}\mathbf{p}\,\frac{[G_{\rm fi}(\mathbf{p}-\mathbf{k}_f)-G_{\rm fi}(\mathbf{q})][|\mathbf{p}-\mathbf{k}_i-\mathbf{k}_f|^2+(\epsilon-ik_i)^2]^{-i\nu}}{|\mathbf{p}-\mathbf{k}_i|^{(2+2i\nu)}|\mathbf{p}-\mathbf{k}_f|^{(2-2i\nu)}(p^2-k_i^2-i\alpha)^{(1-i\nu)}}. \tag{12}$$

If we put $\nu = 0$, (11) and (12) reduce respectively to

$$T_{\rm B2}^N = \lim_{\epsilon\to 0}\int \mathrm{d}\mathbf{p}\,\frac{[G_{\rm fi}(\mathbf{k}_i-\mathbf{p})-G_{\rm fi}(0)]}{|\mathbf{p}-\mathbf{k}_i|^2|\mathbf{p}-\mathbf{k}_f|^2(p^2-k_i^2-i\epsilon)}, \tag{13}$$

and

$$T_{\rm B2}^E = \lim_{\alpha\to 0}\int \mathrm{d}\mathbf{p}\,\frac{[G_{\rm fi}(\mathbf{p}-\mathbf{k}_f)-G_{\rm fi}(\mathbf{q})]}{|\mathbf{p}-\mathbf{k}_i|^2|\mathbf{p}-\mathbf{k}_f|^2(p^2-k_i^2-i\alpha)}. \tag{14}$$

We may verify that, apart from a constant factor, (13) and (14) are the second Born terms for the interactions $+1/r_0$, and $+1/r_{01}$, respectively, calculated in the closure approximation in which the atomic energies of all intermediate states are taken to be ϵ_i, the energy of the initial atomic state [15, 16]. Thus, the TCW model fully includes the closure approximation second Born (CSB2) terms. It is also clear from (8), (10)–(12) that if the TCW amplitude $T_{\rm fi}$, eq. (7), is (formally) expanded in powers of ν, the term containing ν^k arises from the $(k+1)$th order Born term. Thus, we see that the (TCW) model contains contributions from both real and imaginary parts of all higher order (≥ 3) Born terms in this closure approximation.

We may note that, in earlier studies of electron impact excitation of hydrogen ([10], [13]), the first order amplitude $T_{\rm fi}^{(1)}$, eq. (8), was obtained by applying a peaking approximation to the VPS amplitude, i.e. the VPS amplitude was approximated by $T_{\rm fi}^{(1)}$. The first order (or peaking approximation) amplitude $T_{\rm fi}^{(1)}$, eq. (8), being an even function of ν, contains contributions from only odd order Born terms. Evidently, it is a poor approximation to the TCW amplitude $T_{\rm fi}$, eq. (7), since it does not include the second Born amplitude and also the effects of higher order perturbation terms contained in $T_{\rm fi}^{(2)}$, eq. (10).

Obviously, the closure approximation inherent in the TCW theory becomes more accurate when the energy difference between the initial and final atomic states is small and, in addition, there are a large number of intermediate atomic states, grouped around the initial and final atomic states, having nearly the same energy as the initial (or final)

atomic state [20]. Since this condition is encountered in transitions between excited states, the TCW model is better suited for excited-state to excited-state transitions. In this context, an appealing feature of the integral expressions (11) and (12) and hence the transition amplitude T_{fi} is that they depend on the initial and final atomic states only through the form factor $G_{fi}(\mathbf{p})$. Since the form factor $G_{fi}(\mathbf{p})$ is available in closed form for arbitrary initial and final states, the integral expression for the transition amplitude T_{fi} is free from parametric differentiation. We may mention that an alternative method for evaluation of the integrals containing two Coulomb wave functions used in the literature (see for example, [1, 13, 19]) depends on the technique of parametric differentiation. Although this method of parametric differentiation can be applied to excitation of $n = 2$ states, such a method, as is well known, becomes very cumbersome for higher excited atomic states. For this reason, we have preferred to develop numerical techniques for directly computing the three-dimensional integrals (11) and (12).

We have numerically evaluated the three-dimensional integrals (11) and (12) in momentum p-space for the $1s \rightarrow 2s$ excitation of hydrogen induced by electron impact by using spherical polar coordinates $\mathbf{p} \equiv (p, \theta_p, \varphi_p)$. We have suitably divided the range $(0 \leq p \leq \infty,\ 0 \leq \theta_p \leq \pi,\ 0 \leq \varphi_p \leq 2\pi)$ of integrals into some number of intervals (depending on the incident energy) and then applied the Gaussian quadrature formulas in each of these intervals. Careful treatment was given to the regions near $\mathbf{p} = \mathbf{k}_i$ and $\mathbf{p} = \mathbf{k}_f$.

We see that the integrands in (11) and (12) contain the Coulomb distorted Green's function $1/(p^2 - k_i^2 - i\epsilon)^{(1\pm i\nu)}$, having the $i\epsilon$ prescription, i.e. the limiting $\epsilon \rightarrow 0$ procedure. A straightforward application of standard numerical quadrature formulas (e.g. the Gaussian formula [9]) to the integrals in (11) and (12) would miss the contributions originating from the $i\epsilon$ prescription. To evaluate accurately the contributions coming from the $i\epsilon$ prescription in the neighbourhood of $p = k_i$ [and also the singular region around $\mathbf{p} \rightarrow \mathbf{k}_i$], we have given special attention to the radial p-integration from $p = k_i - \Delta_1$ to $p = k_i + \Delta_1$, where Δ_1 is some small positive number and devised an auxiliary integral. The nature of the integrand becomes further involved due to the presence of the factors $[G_{fi}(\mathbf{k}_i - \mathbf{p}) - G_{fi}(0)]/|\mathbf{p} - \mathbf{k}_i|^{(2-2i\nu)}$ and $[G_{fi}(\mathbf{p} - \mathbf{k}_f) - G_{fi}(\mathbf{q})]/|\mathbf{p} - \mathbf{k}_i|^{(2+2i\nu)}$ in (11) and (12), respectively, since these terms become singular as $\mathbf{p} \rightarrow \mathbf{k}_i$. To deal with the $\mathbf{p} \sim \mathbf{k}_i$ region, we have expanded the form factors $G_{fi}(\mathbf{p})$ for the $1s \rightarrow 2s$ excitation in powers of $(\mathbf{p} - \mathbf{k}_i)$ and retained the leading term. We have used the leading term to devise auxiliary integrals. In the region $\mathbf{p} \sim \mathbf{k}_i$, the integrands of the auxiliary integrals are the limiting forms of the corresponding integrands of the actual integrals (11) and (12). Moreover, the auxiliary integrals are such that they can be either evaluated analytically or accurately by numerical quadrature. Therefore, the difference of the actual integrands and the auxiliary integrands is smooth in this region and can be evaluated by numerical quadrature. The complete procedure for the $1s \rightarrow 2s$ excitation is rather lengthy and cumbersome and therefore will not be described here. It may be mentioned that the form of the auxiliary integrals depends sensitively on the form of $G_{fi}(\mathbf{p})$ and hence is different for different initial and final states. The prescription for numerical integration near this region $(\mathbf{p} \sim \mathbf{k}_i)$ for transitions involving non-spherical states (for example, $1s \rightarrow 2p$ excitation) becomes much more involved and we are working on this problem.

We also see that because of the factors $[|\mathbf{p} - \mathbf{k}_i - \mathbf{k}_f|^2 + (\alpha - ik_i)^2]^{+i\nu}/|\mathbf{p} - \mathbf{k}_f|^{(2+2i\nu)}$ and $[|\mathbf{p} - \mathbf{k}_i - \mathbf{k}_f|^2 + (\epsilon - ik_i)^2]^{-i\nu}/|\mathbf{p} - \mathbf{k}_f|^{(2-2i\nu)}$ in (11) and (12), respectively, the

ntegrands become singular as $\mathbf{p} \to \mathbf{k}_f$. However, the integrals (over $\mathbf{p}$) containing these actors in the neighbourhood of $\mathbf{p} = \mathbf{k}_f$ are well defined and can be carried out nalytically. The method for performing these integrals analytically is rather cumbersome nd lengthy and we shall not discuss it here. Our computer program includes these esults.

We have tested the computer program by devising some test integrals which can be valuated analytically and whose integrands in the neighbourhood of $\mathbf{p} = \mathbf{k}_i$ and $\mathbf{p} = \mathbf{k}_f$ ave a very similar behaviour to those of (11) and (12). The program reproduces the nalytical results to desired numerical accuracy.

. Results and discussions

Ve have computed the cross sections of the TCW and closure approximation second Born CSB2) models for the $1s \to 2s$ excitation of hydrogen by electrons in the energy range of 4.4 to 400 eV and have taken into account the exchange contributions by using the)chkur approximation [15]. The computation of the differential cross section at one cattering angle for a single incident energy requires about 30 minutes of CPU time in the BM 6000/R machine.

For the $n = 2$ excitation of hydrogen, a large number of sophisticated perturbative and onperturbative theories exists [15, 16, 21–24] which include the close coupling (40CC) nethod based on a large basis set containing pseudostates of Wyngaarden and Walters 25], the convergent close-coupling (CCC) theory of Bray and Stelbovics [26] and Bubelev *et al* [27], the coupled-channel optical (CCO) model of Bray *et al* [28], the ntermediate R-matrix method (IERM) of Scholtz *et al* [29], the exact second-order istorted wave (DWB2) theory of Madison *et al* [30] and Bubelev *et al* [27]. As stated by Bray and Stelbovics [26], among the nonperturbative theories, the results of the onvergent close-coupling (CCC) theory, the coupled-channel optical (CCO) model of Bray *et al* [28], the intermediate R-matrix method (IERM) of Scholtz *et al* [29] and the lose-coupling (40CC) theory of Wyngaarden and Walters [25] are in excellent agreement vith each other at intermediate energies. Wyngaarden and Walters [25] have also resented the 40CC results in tabular forms. Therefore, we have compared the TCW esults with the 40CC results and, for the sake of clarity, we have not included in the igures the results of the CCC, CCO, and IERM theories.

.1 *Electron impact excitation of the 2s state at 54.4 eV*

For the electron impact excitation of the ground $1s$ state of the hydrogen to the $2s$ state, he incident energy 54.4 eV is special, since only at this energy, the differential cross ections have been measured by Williams in 1981 [31] (see also Williams and Wang [32] or a review of the $n = 3$ excitation of hydrogen). Figure 1 compares the differential cross ections (shown as the solid curve) of the TCW model with the experimetal data. It is een that the TCW results (with the exchange effects included) are in very good greement with the experimental data for scattering angles $\Theta \leq 100°$ but they are smaller han the experimental results for larger scattering angles $\Theta > 100°$. This figure also lisplays the results of the TCW model calculated by ignoring the exchange effects

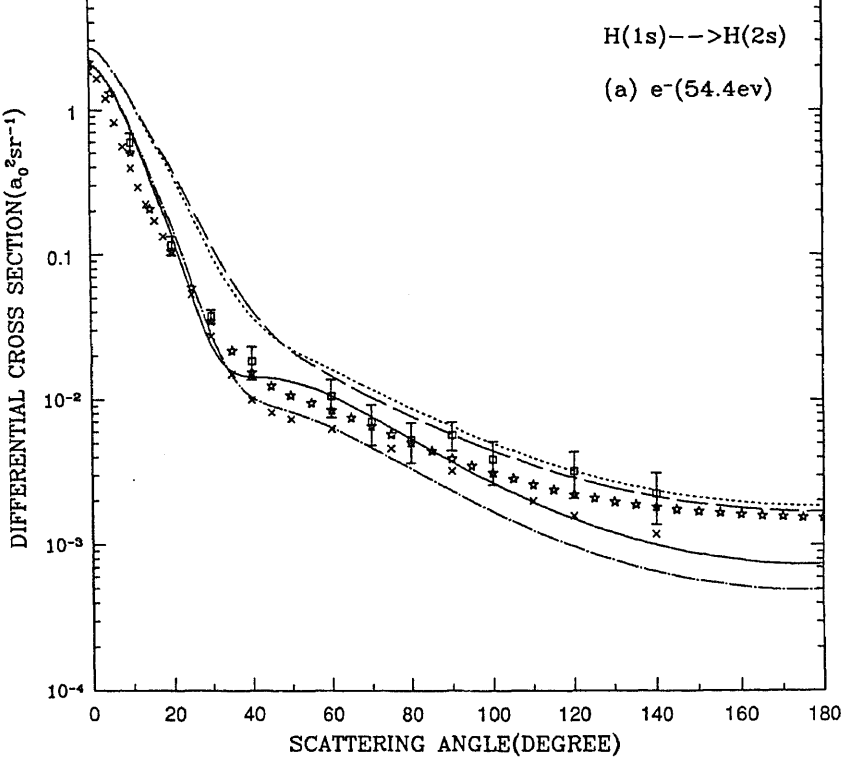


Figure 1. Differential cross sections for $e^- + \mathrm{H}(1s) \longrightarrow e^- + \mathrm{H}(2s)$ at 54.4 eV. Solid line: TCW results with exchange; dash-dot line: TCW results without exchange; dottted line: CSB2 results with exchange; dashed line: CSB2 results without exchange; stars: DWB2 results [27, 30]; crosses: 40CC results [25]; open squares with error bars: experimental results of Williams [31].

(shown as dash-dotted line). We see that the exchange effects are important at this energy. It is also seen that the CSB2 results deviate significantly from the TCW cross sections showing the importance of the higher order ($n \geq 3$) Born terms included in the TCW theory. The differential cross sections of the first Born theory, and the peaking approximation given by $T_{\mathrm{fi}}^{(1)}$, eq. (8) are not shown in the figure. The peaking approximation predicts that the differential cross sections have a shape similar to that given by the first Born theory. It is well documented in the literature (see e.g. [16]) that the first Born gives a poor description of experiment.

Figure 1 also displays the theoretical results (shown as crosses) of Wyngaarden and Walters [25] calculated by using the close coupling (40CC) method based on a large basis set consisting of s, p and d states which also includes pseudostates to account for the couplings to the continuum (and the bound states not explicitly included in the basis set). In this 40CC method, couplings to the target states of higher angular momenta ($l > 2$) are treated perturbatively in the second Born approximation. The TCW theory predicts a

ightly broader peak for smaller scattering angles in comparison to the 40CC theory of /yngaarden and Walters [25] and gives somewhat better agreement with the xperimental data in this region. Except for small scattering angles (given below), the sults of the TCW model are in good overall agreement with those of 40CC method. An teresting feature of the TCW and 40CC (and also the CCC [26, 27]) theories is that they hibit a broad shoulder around $\Theta \approx 40°$.

Madison *et al* [30] and Bubelev *et al* [27] have performed calculations of the oss sections for the $n = 2$ excitation of hydrogen by electron impact in the energy range)–200 eV in a distorted wave second order (DWB2) theory which includes exactly the cond order direct and exchange amplitudes. The DWB2 results are shown (as stars) in gure 1. We see that the TCW model (the solid curve) is in good harmony with the WB2 theory (the stars) for scattering angles $\Theta \leq 100°$. For larger scattering angles $\Theta > 100°$), the differential cross sections of the DWB2 theory tend towards the results of e closure approximation second Born (CSB2) model, calculated by using (13) and (14) hown as the dotted curve). It is interesting to note that the differential cross sections of e CSB2 and DWB2 approximations agree remarkably well with the experimental data r $\Theta \geq 90°$, whereas for smaller scattering angles, the CSB2 approximation gives a poor

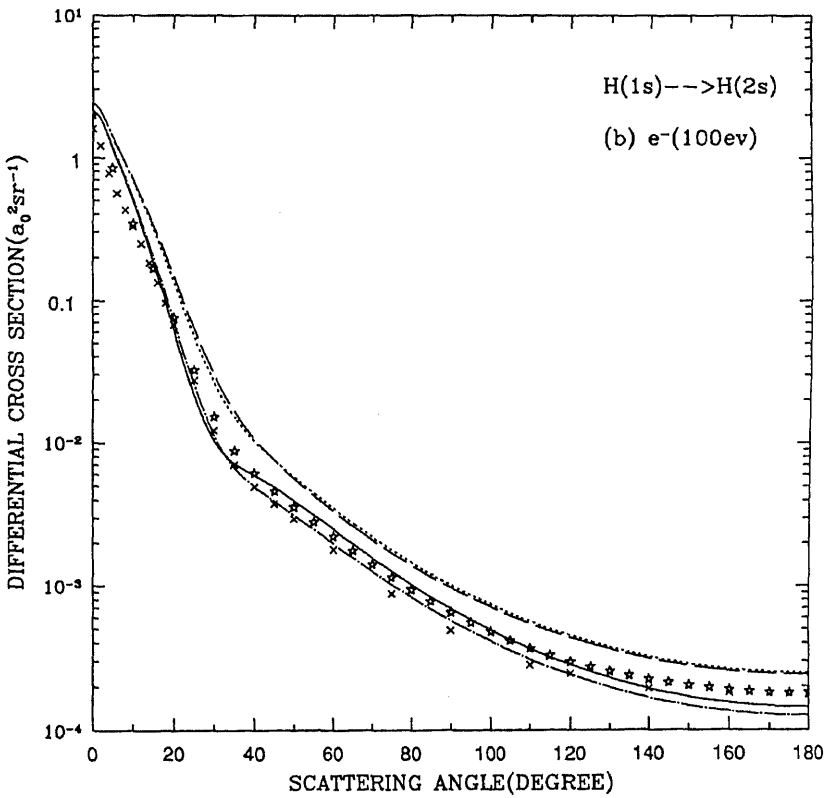


Figure 2. Differential cross sections for $e^- + H(1s) \longrightarrow e^- + H(2s)$ at 100 eV. The descriptions of the curves are the same as in figure 1.

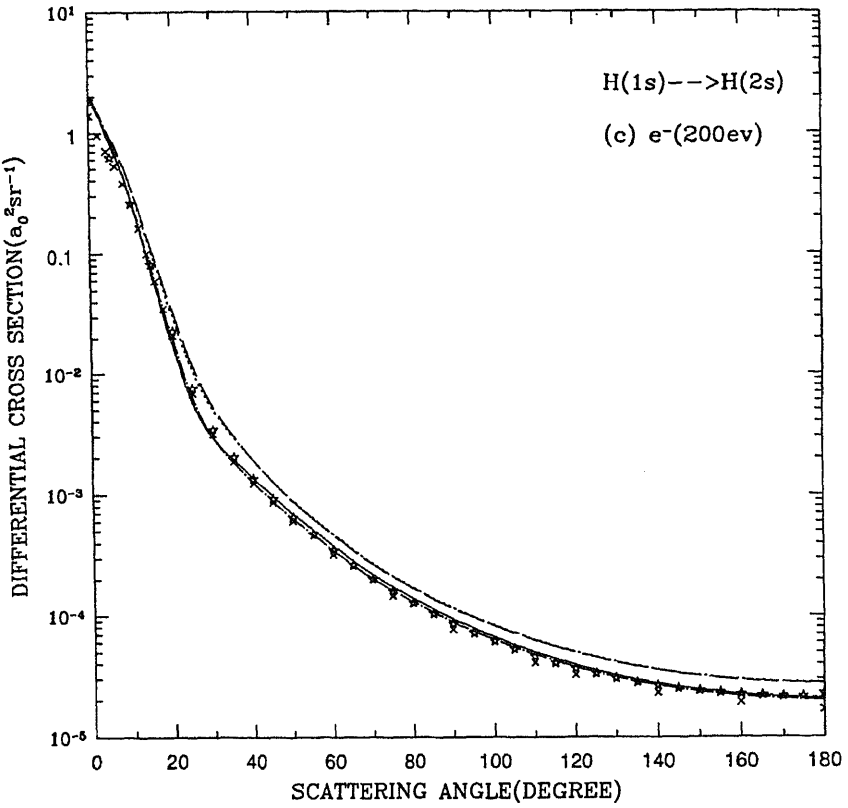


Figure 3. Differential cross sections for $e^- + \mathrm{H}(1s) \longrightarrow e^- + \mathrm{H}(2s)$ at 200 eV. The descriptions of the curves are the same as in figure 1.

description of the experimental data. Another interesting feature to note is that in contrast to the TCW, 40CC and CCC theories, the DWB2 and CSB2 models do not predict a broad shoulder around $\Theta \approx 40°$.

3.2 *Electron impact excitation of the* 2*s state for* 100–400 *eV*

The differential cross sections calculated in the TCW model (the solid curve) and the CSB2 approximation [both with (the long dashed curve) and without (the dotted curve) exchange contributions] are shown in figures 2–5 for incident electron energies 100–400 eV. For these energies, there are no experimental data. The results of the 40CC approximation of Wyngaarden and Walters [25] are available for 100, 200 and 300 eV which are also shown as the crosses. The stars in figures 2 and 3 represent the results of the DWB2 theory calculated by Madison *et al* [30] and Bubelev *et al* [27]. We see from these figures that for incident energy $\geq$ 200 eV, the exchange contribution is very small. The TCW model predicts somewhat larger values of the differential cross sections in the forward direction in comparison to the 40CC and the DWB2 models. This feature is due to the average excitation energy being equal to ϵ_i, the energy of the initial atomic state.

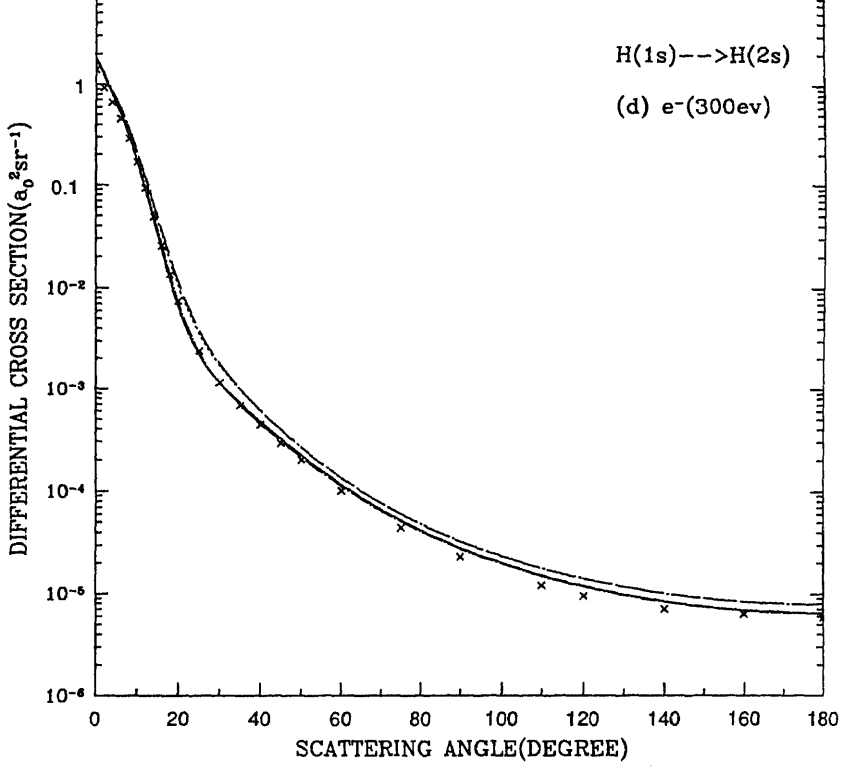


Figure 4. Differential cross sections for $e^- + \mathrm{H}(1s) \longrightarrow e^- + \mathrm{H}(2s)$ at 300 eV. The descriptions of the curves are the same as in figure 1.

'e see that for the scattering angles $\Theta > 20°$, there is good overall agreement among the 'edictions of the TCW, 40CC and DWB2 methods.

3 *Integrated cross sections*

'e have computed the integrated cross sections of the TCW, CSB2 and first orn theories and also of the peaking approximation given by $T_{\mathrm{fi}}^{(1)}$, eq. (8). In our ılculations, we have included the exchange contribution by using the Ochkur ɔproximation. Table 1 compares the integrated cross sections calculated in various ıodels. The importance of the higher order Born terms included in the TCW theory ıce again becomes clear from the fact that the integrated CSB2 cross sections ffer significantly from the integrated TCW cross sections in the energy range under ɔnsideration. Indeed, the CSB2 theory predicts the largest values of the integrated oss sections. The peaking approximation also gives larger integrated cross sections lower energies showing the importance of the remaining terms included in the CW theory. From the table we also see that the TCW model predicts larger values of e integrated cross section in comparison to the 40CC close coupling results of

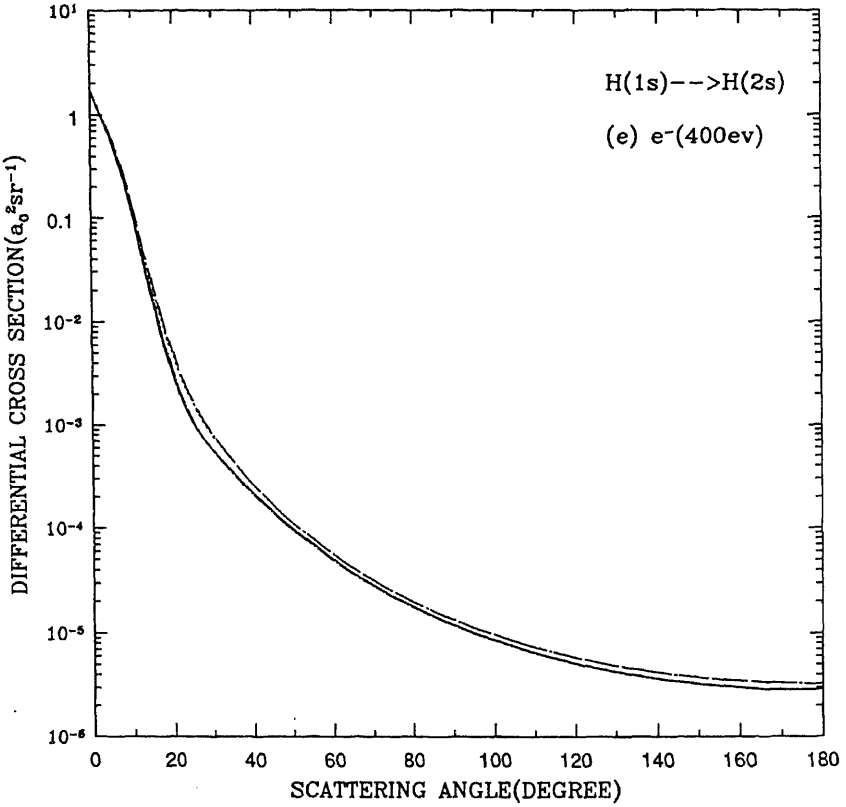


Figure 5. Differential cross sections for $e^- + \mathrm{H}(1s) \longrightarrow e^- + \mathrm{H}(2s)$ at 400 eV. The descriptions of the curves are the same as in figure 1.

Table 1. The integrated cross sections (in units of a_0^2) as a function of the incident electron energy for the $1s \longrightarrow 2s$ excitation of atomic hydrogen by electron impact. The integrated cross sections of the TCW, CSB2, peaking approximation and first Born theories include the exchange contributions through the Ochkur approximation.

Electron energy	54.4 eV	100 eV	200 eV	300 eV	400 eV
TCW	0.298	0.174	0.091	0.062	0.046
CSB2	0.568	0.261	0.114	0.072	0.052
40CC	0.204	0.126	0.079	0.057	—
Peaking approx.	0.437	0.205	0.096	0.063	0.047
First Born	0.289	0.170	0.090	0.061	0.046

Wyngaarden and Walters [25]. This is due to the larger values of the differential cross section in the forward directions (see figures 1–4) resulting from the closure approximation imbedded in the TCW model. But, as the energy increases, the integrated cross sections of both the TCW and 40CC models approach the first Born values, as one would expect.

4. Concluding remarks

In the context of recent interest in the BBK wave function for ionization, we have discussed the two Coulomb waves (TCW) model of Dewangan [8] for direct excitation. We have seen that the TCW model for excitation corresponds to the BBK model for ionization. To examine the predictions of the TCW model, we have studied the $1s \rightarrow 2s$ excitation of hydrogen by electron impact in the energy range of 54.4 to 400 eV and compared them with the results of other sophisticated perturbative and nonperturbative models. We have seen that, because of the large difference in the eigenenergies of the $1s$ and $2s$ states, the closure approximation inherent in the model may not be well justified for the $1s \rightarrow 2s$ excitation as for transitions between excited states. In spite of this limitation, the TCW results give good agreement with the experimental differential cross sections at 54.4 eV by electron impact and also compare reasonably well with the predictions of other current sophisticated theoretical models.

References

[1] M Brauner, J S Briggs and H Klar, *J. Phys.* **B22**, 2265–2287 (1989)
[2] C R Garibotti and J E Miraglia, *Phys. Rev.* **A21**, 572–580 (1980)
[3] A Franz and P L Altick, *J. Phys.* **B25**, 1557 (1992)
[4] J Berakdar and J S Briggs, *Phys. Rev. Lett.* **72**, 3799 (1994)
[5] J Berakdar, A Engelns and H Klar, *J. Phys.* **B72**, 1109 (1996)
[6] F Maulbetsch and J S Briggs, *J. Phys.* **B26**, L647–L652 (1993)
[7] F Maulbetsch and J S Briggs, *J. Phys.* **B26**, 1679–1696 (1993)
[8] D P Dewangan, *J. Phys.* **B16**, L595–99 (1983)
[9] M Abramowitz and I A Stegun, *Handbook of mathematical functions* (Dover, New York, 1968)
[10] L Vainshtein, L Presnyakov and I Sobelman, *Sov. Phys.–JETP* **18**, 1383 (1964)
[11] I M Cheshire, *Proc. Phys. Soc.* **84**, 89 (1964)
[12] D P Dewangan and B H Bransden, *J. Phys.* **B15**, 4561–76 (1982)
[13] C Sinha, K Roy and N C Sil, *J. Phys.* **B20**, 2545–2552 (1987)
[14] D P Dewangan and S B Khadkikar, *J. Phys.* **B18**, L751–54 (1985)
[15] F W Byron Jr. and C J Joachain, *Phys. Rep.* **34**, 233–324 (1977)
[16] H R J Walters, *Phys. Rep.* **116**, 1 (1984)
[17] D P Dewangan and J Eichler, *Phys. Rep.* **247**, 59 (1994)
[18] B H Bransden and M R C McDowell, *Charge exchange and the theory of ion-atom collisions* (Oxford, 1992)
[19] X J Chen, B Wen, Y W Xu and Y Y Zheng, *J. Phys.* **B25**, 4661–72 (1992)
[20] D P Dewangan and H S Chakraborty, *J. Phys.* **B24**, L263–268 (1991)
[21] J Callaway, *Adv. Phys.* **29**, 771 (1980)
[22] J Callaway and K Unnikrishnan, *Phys. Rev.* **A48**, 4292–98 (1993)
[23] Y D Yang, J Callaway and K Unnikrishnan, *Phys. Rev.* **A49**, 1854 (1994)
[24] I Bray and A T Stelbovics, *Adv. At. Mol. Opt. Phys.* **35**, 209 (1995)
[25] W L van Wyngaarden and H R J Walters, *J. Phys.* **B19**, 1827 (1986)
[26] I Bray and A T Stelbovics, *Phys. Rev.* **A46**, 6995–7011 (1992)
[27] V E Bubelev, D H Madison, I Bray and A T Stelbovics, *J. Phys.* **B28**, 4619–4638 (1995)
[28] I Bray, D A Konovalov and I E McCarthy, *Phys. Rev.* **A44**, 5586 (1991)
[29] T T Scholz, H R J Walters, P G Burke and M P Scott, *J. Phys.* **B24**, 2097 (1991)
[30] D H Madison, I Bray and I E McCarthy, *J. Phys.* **B24**, 3861–3888 (1991)
[31] J F Williams, *J. Phys.* **B14**, 1197–217 (1981)
[32] J F Williams and J B Wang, *Adv. At. Mol. Opt. Phys.* **32**, 407 (1994)

—Journal of
physics

November 1997
pp. 493–504

On the water-induced critical double point in a polymer solution: the role of isotopic substitution

B M JAFFAR ALI, S VENKATACHALAM* and A KUMAR†
Department of Physics, Indian Institute of Science, Bangalore 560 012, India
* Present address: Department of Physics, University of California, San Diego, La Jolla, CA 92093, USA
† Corresponding author.
E-mail phyanil@physics.iisc.ernet.in

MS received 14 July 1997; revised 22 October 1997

Abstract. This paper examines the effect of substitution of water by heavy water in a polymer solution of polystyrene (molecular weight = 13000) and acetone. A critical double point (CDP), at which the upper and the lower partially-miscible regions merge, occurs at nearly the same coordinates as for the system [polystyrene + acetone + water]. The shape of the critical line for [polystyrene + acetone + heavy water] is highly asymmetric. An explanation for the occurrence of the water-induced CDP in [polystyrene + acetone] is advanced in terms of the interplay between contact energy dissimilarity and free-volume disparity of the polymer and the solvent. The question of the possible existence of a one-phase *hole* in an hourglass phase diagram is addressed in [polystyrene + acetone + water]. Our data exclude such a possibility.

Keywords. Liquid–liquid transitions; critical double point; polymer demixing; isotopic substitution.

PACS Nos 64.70; 64.60

1. Introduction

Polymer solutions exhibit a few novel thermodynamic states, e.g., theta point, critical double point (CDP) [1, 2] and tricritical point [3] induced by variations in the molecular weight of the polymer or the applied pressure. Polymer solutions are known to share the Ising universality class with binary liquid mixtures [1, 2]. Polymer solutions show a high temperature partially-miscible region that terminates at lower critical solution temperature, T_L and a low temperature partially-miscible region that terminates at upper critical solution temperature, T_U [4–6]. The CDP is defined as a point where T_L and T_U are made to merge, i.e., the miscibility gap $\Delta T[= |T_L - T_U|]$ becomes zero. Other examples of systems that display CDP are aqueous alcohols [secondary butyl alcohol + tertiary butyl alcohol + water] [7], binary gases (Ne-Kr) [8] and mixtures of liquid sulphur and hydrocarbons (e.g., benzene or toluene). A double critical point, DCP (where a closed-loop phase diagram vanishes) is identical to a CDP except for some finer details concerning the approach to these points [6].

In practice, the role of pressure in multicomponent liquid systems as well as polymer solutions is simulated by isotopic/isomeric substitutions [6, 7, 9] or by additional

understood within the framework of the Prigogine solution theory as modified by Patterson [4, 5, 10]. The quantity of interest here is the interaction parameter, χ, which reflects the interplay of two competing terms, viz., the contact energy dissimilarity and free-volume disparity of the polymer and the solvent. The expression for χ, as given by the Patterson and the Flory theories [4], is

$$\chi = \frac{-U_1}{RT}\nu^2 + \left(\frac{C_{p1}}{2R}\right)\tau^2 \tag{1}$$

where U_1 is the configurational energy of the solvent, C_{p1}, the first derivative of U_1 with respect to T, is the configurational heat capacity of the solvent, and R is the gas constant. The parameter ν^2 expresses the *difference* of the cohesive energy and size between the solvent molecule and the polymer segment. The parameter τ reflects the free-volume change, which occurs due to large *differences* in the thermal expansion coefficients of the polymer and the solvent.

The extent of miscibility or the extent of stability of the mixed state is determined by the Flory theory from the combinatorial part of ΔG_M, viz., the total free energy of mixing. From this we get an expression for the critical value of χ, χ_c, as follows [4, 10]

$$\chi_c(T,P) = \frac{1}{2}\left[1 + \left(\frac{v_1}{v_2}\right)^{1/2}\right]^2, \tag{2}$$

where v_1 and v_2 are the molar volumes of the solvent and the polymer, respectively. It should be noted that χ_c decreases with an increase in molecular weight of the polymer.

The first term in (1) usually decreases with temperature (T), whereas the second term is an increasing function of temperature. Typically, a $\chi - T$ curve (figure 1) has a minimum

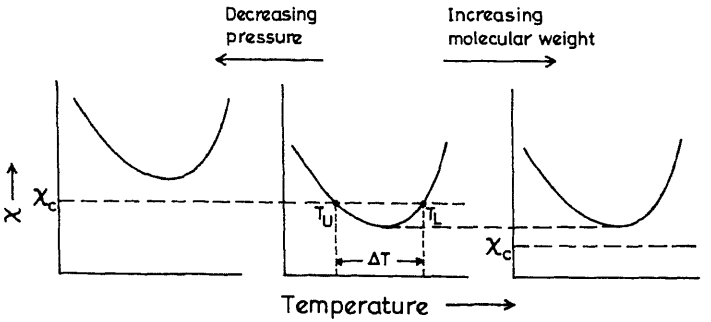


Figure 1. The behaviour of the $\chi - T$ curve, and χ_c with change in pressure and molecular weight of the polymer. These changes lead to an alteration in the miscibility behaviour (i.e., ΔT value) and the realization of a critical double point ($\Delta T = 0$) as well as appearance of an *hourglass* phase diagram.

n χ (i.e., $\partial\chi/\partial T \to 0$ or $\partial T/\partial\chi \to \infty$). A *finite* ΔT is attained when the line corresponding to χ_c intersects this curve at temperatures T_U and T_L. The difference between these two temperatures, $\Delta T[= |T_L - T_U|]$, denotes the size of the miscibility gap.

Increasing the molecular weight of the polymer without changing the nature of the $\chi - T$ curve leads to a decrease in the χ_c value (figure 1). Hence, at some value of molecular weight, the χ_c line can become *tangential* to the $\chi - T$ curve, leading to the occurrence of a CDP ($\Delta T = 0$). The situation in which the $\chi - T$ curve lies above the χ_c value corresponds to the *hourglass* phase diagram.

Alternatively, a decrease in pressure (P) can *increase* the free-volume differences between the two components and thus shift the $\chi - T$ curve *upwards* (figure 1). However, the χ_c line does not vary with pressure. Thus a condition of χ_c line being *tangential* to the $\chi - T$ curve is again realized, resulting in a CDP. Any further decrease in pressure yields an *hourglass* phase diagram. In this manner, one can explain the occurrence of a CDP in a polymer solution in two different ways.

Changing the molecular weight of polystyrene from 4800 to 19800 in [PS + AC] leads to a CDP at a molecular weight of 19800 and a temperature of $\simeq 70°C$ [4]. By variation of pressure in [PS (molecular weight 20000) + AC], the CDP is realized [2,5] at a pressure of 3.33 bar and a temperature of 69.86°C. We have shown earlier [1] that a small quantity of the third component, i.e., water, can induce a CDP in [PS (molecular weight 13000) + AC] at 89.258°C for 0.007 weight fraction of water.

The objectives of this work are: (a) to discern the effect of isotopic (H/D) substitution on the nature of CDP in [PS + AC] solutions, i.e., to impart generality to our earlier observation of CDP in this system, and (b) to uncover finer details of *hourglass* phase diagram in close neighbourhood of CDP – in the light of the observation [11] of a one-phase *hole* or *loop* inside the *hourglass* configuration.

2. Experiment

Samples were made using polystyrene (PS) (Pressure Chemical Co.) of molecular weight 13000 and polydispersity of less than 6%, i.e., $M_w/M_n = 1.06$, where M_w is weight average molecular weight and M_n is number average molecular weight. The acetone (AC) was of spectroscopic grade (99.8%, Ranbaxy, India). Heavy water was of isotopic purity 99.6% (BARC, India). Water used was triple distilled in an all-quartz distiller.

The miscibility range, $\Delta T(= T_L - T_U)$, is controlled by the weight fraction (W) of water or heavy water in the polymer solution [PS + AC]. We believe that W simulates the condition of a thermodynamic field in the sense that the nature of W–T graph is similar to that of a P–T plot for this system [2]. The amount of water/heavy water involved is extremely small, varying from 0.0008 to 0.0080 weight fraction of the total mixture. To permit the addition of such small quantities, a stringent sample preparation procedure was adopted.

The critical concentration (w_c, weight fraction of polystyrene) for [PS + AC + heavy water] system was determined and it was found to be nearly constant at 0.23 for all ΔTs, i.e., w_c had no ΔT dependence. So, to realize the critical line correctly, it was necessary to maintain the constancy of the critical concentration (w_c) for samples with varying ΔTs. Stock solutions of [AC + heavy water] were prepared in the required concentrations

corresponding to different values of W and concomitant with a fixed amount of PS (w_c) under similar conditions. Requisite amounts of this solution were added to the correctly weighed amounts of PS. A stock solution of [AC + heavy water] was made so as to overcome the problem of actual addition of extremely small quantities of heavy water ($\simeq 0.0005$ g) to the [PS + AC] system.

In the case of the determination of *hourglass* curve in [PS + AC + water] system, the quantity W (weight fraction of water in the mixture) is to be kept constant. A stock solution of [AC + water] was prepared. However, the quantities of AC and water added were in such proportions that they corresponded to a weight fraction (w) of polystyrene $= 0.5$. A sample with $w = 0.5$ was prepared from this stock solution. Subsequently, several samples with different weight fractions of polystyrene ($w < 0.5$) were prepared for a fixed W. The total amount of water and acetone required for each of these samples was calculated. An amount of stock solution containing the required amount of water was then added. The requisite amount of additional acetone was calculated and added separately since the stock solution of [AC + water] was prepared for $w = 0.5$. Once again, the difficulty of having to add exceptionally small quantities of water was circumvented by using the stock solution of [AC + water]. This method

Table 1. Details of the samples for [polystyrene + acetone + heavy water] used to study the shape of the critical line (figures 2–4). W denotes the weight fraction of heavy water and w_c signifies the critical weight fraction of polystyrene. T_U and T_L indicate the upper and lower consolute points, respectively.

W ±0.0001	w_c ±0.003	T_U(°C) ±0.03	T_L(°C)	ΔT(°C) ($T_L - T_U$)
0.0020	0.23	33.10	118.8 ±1	85.7 ±1.03
0.0030	0.23	38.730	117.0 ±1	78.27 ±1.03
0.0040	0.23	46.264	115.0 ±1	68.736 ±1.03
0.0050	0.23	53.75	112.5 ±0.4	58.75 ±0.43
0.0060	0.23	62.247	109.4 ±0.1	47.153 ±0.13
0.0065	0.23	66.782	107.5 ±0.1	40.718 ±0.13
0.0070	0.23	70.824	104.6 ±0.1	33.776 ±0.13
0.0075	0.23	77.315	101.6 ±0.1	24.285 ±0.13
0.0078	0.23	83.802	98.6 ±0.05	14.798 ±0.08
0.0080	0.23	88.68	95.725 ±0.05	7.045 ±0.08
0.00805	0.23	90.665	94.115 ±0.05	3.45 ±0.08

Table 2. The coexistence data for the *hourglass* coexistence curve shown in figure 6. w denotes the weight fraction of polystyrene.

T(°C)	w^+ ±0.003	w^- ±0.003	Δw $(w^+ - w^-)$
	$W = 0.0072$		
103.5 ±0.1	0.45	0.158	0.292
101.2 ±0.1	0.436	0.17	0.266
97.523 ±0.05	0.40	0.19	0.210
94.710 ±0.05	0.36	0.197	0.163
94.202 ±0.05	0.352	0.20	0.152
92.730 ±0.03	0.34	0.204	0.136
91.421 ±0.03	0.31	0.206	0.104
90.242 ±0.03	0.281	0.21	0.071
88.420 ±0.03	0.302	0.21	0.092
87.82 ±0.03	0.31	0.209	0.101
85.465 ±0.03	0.34	0.208	0.132
83.76 ±0.03	0.36	0.204	0.156
83.424 ±0.03	0.20	0.36	0.160
80.545 ±0.03	0.40	0.19	0.210
77.220 ±0.03	0.45	0.173	0.277
76.835 ±0.03	0.454	0.17	0.284
70.865 ±0.03	...	0.14	

allowed for a complete control over W and w as well as a constancy of conditions under which different samples were prepared. A typical sample (in all preparations) was about 0.15–0.20 ml in volume. This typically contained about 0.0005 g of water.

Samples were prepared in pyrex cells with Teflon high vacuum stopcocks (Rotaflo). A serious problem encountered in generating the coexistence curve and the critical line is the drift of ΔT, particularly for its small values. This difficulty stems from the acute steepness of the coexistence surface in the vicinity of the CDP [1, 2, 6–8]. Neither

degassing to remove trapped atmospheric oxygen nor preparation of the sample under dry nitrogen improved the temporal stability of the samples for any ΔT. Because only very small quantities of water or heavy water are involved in diminishing the ΔT (tables 1 and 2), a great deal of caution has to be exercised in maintaining the constancy of W. Reliable data could be obtained only if phase separation temperatures were determined soon after the samples were made.

Phase separation temperatures were determined visually [1, 6] by monitoring the onset of intense opalescence. Criticality of the samples was ascertained by the equal-volume criterion. For off-critical samples, phase separation temperatures were marked by the actual appearance of meniscus. The visual measurements were done in a well-stirred paraffin thermostat with a temperature stability better than ± 1 mK in the temperature range 35–95°C. The temperature is measured to an absolute accuracy of ± 60 mK with a platinum resistance thermometer in conjunction with a resistance bridge (Model F17, ASL, UK). The determination of phase separation temperature in the range 95–110°C was carried out in a polyethylene glycol thermostat with a temperature stability of ± 50 mK. For temperature above 110°C, a polyethylene glycol thermostat with a temperature stability of ± 0.5 K was used. It was essential to stir the sample during the course of detection of phase separation temperature. A typical run to determine the upper critical solution temperature (T_U) lasted 6–8 hours and the temperature range scanned was 2°C. Determination of lower critical solution temperature (T_L) lasted 4 hours and the temperature range scanned was about 2°C.

The line of critical points was determined in a single set of experiments. It was observed that the temperature range of persistence of critical opalescence is smaller for [PS + AC + heavy water] than for [PS + AC + water]. Table 1 gives the details of the determination of critical line for [PS + AC + heavy water] and table 2 provides the data for the *hourglass* coexistence curve for [PS + AC + water].

3. Results

The data for T_U and T_L [(PS + AC + heavy water), table 1] were fitted to the following expression

$$(T - T_{CDP}) = A_1(W - W_{CDP})^{A_2} \tag{3}$$

where T_{CDP} and W_{CDP} are the coordinates of the CDP. The data were analysed using a modified form of the program CURFIT [1]. The fit was deemed to be good if the minimum of reduced χ^2 (χ^2_ν) was close to 1 and resulting residuals had random distribution. The best-fit values of the parameters are

$$A_1 = 321.02 \pm 4.59; \quad A_2 = 0.6046 \pm 0.0080; \quad \chi^2_\nu = 1.213$$

and

$$A'_1 = 93.40 \pm 0.54; \quad A'_2 = 0.4380 \pm 0.0018; \quad \chi^2_\nu = 2.220,$$

where A_1 and A_2 refer to the lower section of the critical curve, and A'_1 and A'_2 pertain to the upper section of the critical curve. The coordinates of CDP are given by $T_{CDP} = 92.275 \pm 0.035$°C and $W_{CDP} = 0.00806 \pm 0.00010$. The solid line in figure 2 is

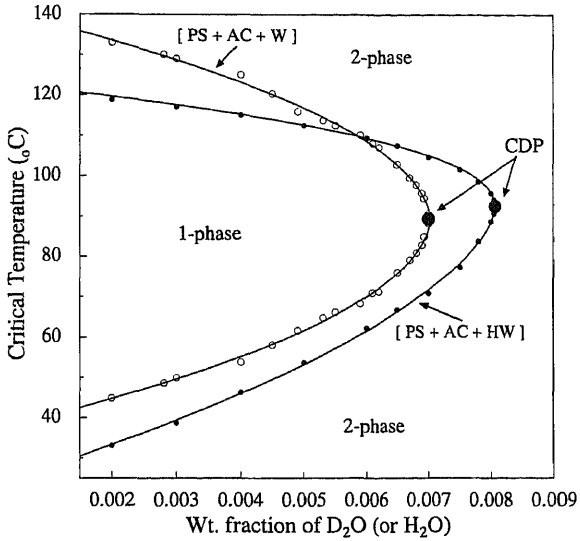


Figure 2. A comparative portrayal of the shape of the critical line for [polystyrene + acetone + water] and for [polystyrene + acetone + heavy water] in terms of *weight fraction* of water or heavy water. Extrema of the curves denote the critical double point (CDP).

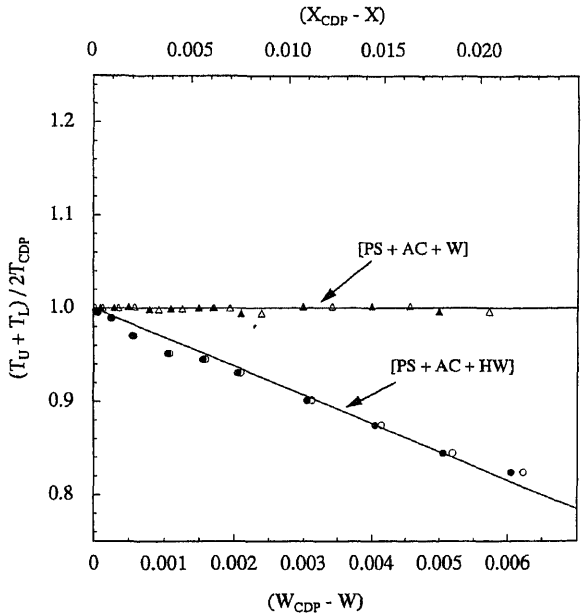


Figure 3. Illustration of the acute asymmetry in the line of critical points for [polystyrene + acetone + heavy water], in contrast to the case for [polystyrene + acetone + water]. The filled symbols refer to the weight fractions (W) of water/heavy water, whereas the corresponding open symbols relate to the mole fraction (X) of water or heavy water.

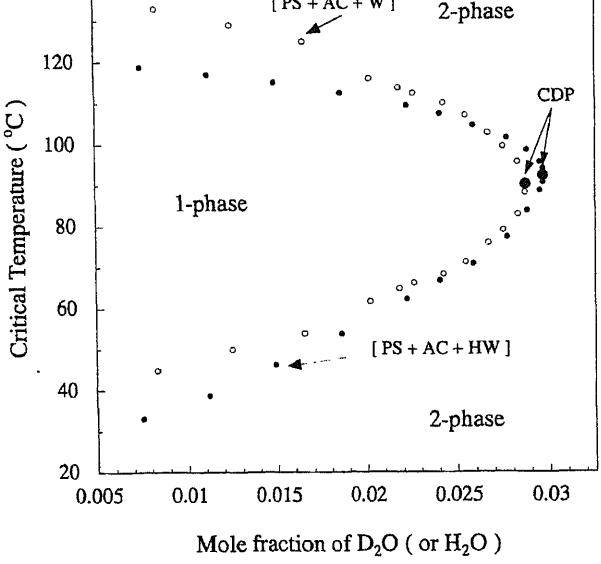


Figure 4. A comparative portrayal of the shape of the critical line for [polystyrene + acetone + water] and for [polystyrene + acetone + heavy water] in terms of *mole fraction* of water or heavy water. Extrema of the curves denote the critical double point (CDP). Note the much closer values of X_{CDP} for the two curves in comparison to the situation depicted in figure 2.

generated by (3). This curve is found to be grossly asymmetric, as reflected in figure 3. If data reduction is confined to the close proximity of CDP, the critical line is fairly symmetric (figure 3). In comparison, the line of critical point for [PS + AC + water] was found to be both symmetric and parabolic [1]. The coordinates of the CDP for [PS + AC + water] are: $T_{CDP} = 89.258 \pm 0.03°C$ and $W_{CDP} = 0.00702 \pm 0.00010$.

In view of the fact that molar volume of water and heavy water are almost identical, the nature of critical lines for [PS + AC + water] and [PS + AC + heavy water] was also examined using mole fractions (X) of water or heavy water. Figure 4 reveals that the two CDPs are much closer in terms of mole fractions, though the asymmetry of the critical line still persists (figures 2 and 3). For instance, the weight fraction ratio of CDPs is 1.15 (figure 2) and the mole fraction ratio of CDPs is virtually unity (1.03), as evidenced by figure 4.

The reentrant critical line in ternary mixture [7] of [secondary butylalcohol + tertiary butylalcohol + water] and in binary gases [8], e.g., (Ne–Kr) are quite asymmetric, particularly away from a CDP.

Although polystyrene (PS) used in the experiments involving [PS + AC + heavy water] was of the same stated purity and polydispersity as the one used [1] in [PS + AC + water], it belonged to a different lot. This fact may have led to the differences in the limiting critical temperatures of the two curves shown in figure 2 for the binary limit of [PS + AC]. However, the ΔT values for [PS + AC] seem to be reasonably closer for the two curves in figure 2.

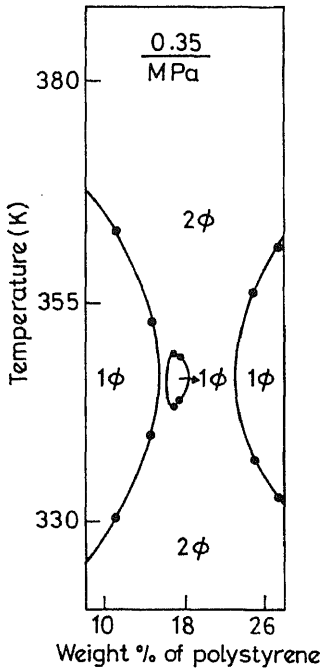


Figure 5. Hourglass coexistence curve for [polystyrene (molecular weight 22400) + acetone] system at 0.35 MPa (3.5 bar). Note the unusual existence of one-phase (1ϕ) *hole* inside the two-phase region (after ref. [11]).

The closest approach to CDP in [PS + AC + heavy water] is signified by $\Delta T = 3.45°C$ (table 1, figure 2). In ternary liquid mixtures, the critical point does not coincide with the apex or the extremum (T_{EL} or T_{EU}) of the relevant coexistence curve [1, 6]. Therefore $\Delta T = 3.45°C$ actually corresponds to the approach to an *effective* CDP to within a few hundreds of mK [1]. For instance, for [PS + AC + water], $\Delta T = 2.34°C$ related to ΔT_E $(= T_{EL} - T_{EU}) = 194$ mK [1].

In a bid to locate the existence of a one-phase (1ϕ) *hole* [11] in an *hourglass* phase diagram (figure 5), we explored the relevant phase boundary for [PS + AC + water]. This *hourglass* phase diagram (figure 6) was determined for $W(= 0.0072) > W_{CDP}$ $(= 0.00702)$ so that it met the condition of being in the close proximity of CDP. As figure 6 illustrates, the entire two-phase (2ϕ) region of the *hourglass* phase diagram is devoid of any one-phase *hole*.

4. Discussion

We had earlier demonstrated [1] that the introduction of very small quantities of water to a polymer solution of [PS + AC] can yield a CDP. The same finding has been reinforced by the substitution of water by heavy water in [PS + AC]. There does not seem to be any profound difference between the CDP in [PS + AC + water] and that in [PS + AC + heavy water] (figure 4). This feature may possibly be due to the fact that minute quantities

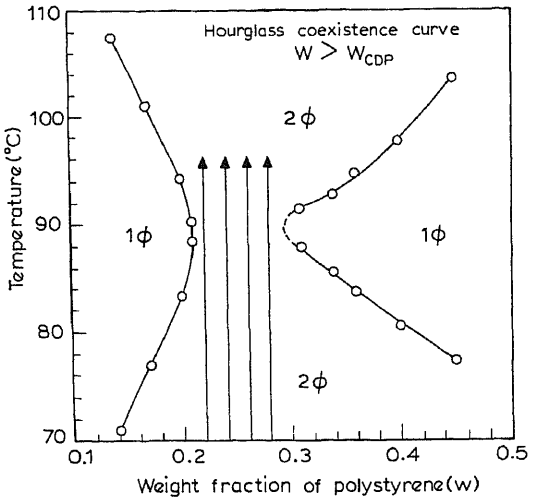


Figure 6. Hourglass coexistence curve for [polystyrene (molecular weight 13000) + acetone + water] for $W(= 0.0072) > W_{CDP}(0.00702)$, where W is the weight fraction of water in the mixture. Arrows indicate the region scanned to uncover the possible existence of a one-phase (1ϕ) *hole* in the vicinity of critical double point, CDP ($w_{CDP} = 0.23$, $T_{CDP} = 89.258°C$).

of either water or heavy water are adequate to realize the merger of upper and lower partially-miscible regions. The earlier experiments concerning isotopic substitution in polymer solution systems had concluded that the deuteration of polymer simulated the applied pressure and the deuteration of the solvent led to simulation of a negative pressure [9].

A striking outcome of H/D substitution in [PS + AC + water] is the asymmetry in the critical curve (figures 2 and 3). It is stressed that in the close proximity of the CDP, both the curves are reasonably symmetric (figure 3). For example, the asymmetry is no more than $\sim 5\%$ for $(W_{CDP} - W) \leq 0.001$. This composition range implies a $\Delta T \leq 24°C$. From the standpoint of the evolution of critical behaviour near CDP, the region of $\Delta T \leq 24°C$ is quite adequate [6]. Besides, Landau–Ginzburg theory and lattice-solution models demand a symmetric, parabolic critical line only in the immediate vicinity of a DCP or a CDP – at least for systems involving small molecules [6]. A possible explanation of the asymmetry of the critical curve for [PS + AC + heavy water], in contrast to that for [PS + AC + water], may be related to the slight differences in the nature of the *exact* $\chi - T$ curves for the two cases. Beyond these considerations, we have no other explanation to account for the asymmetry of the critical curve of [PS + AC + heavy water] in the region *away* from the CDP.

The role of water or heavy water in suppressing ΔT in [PS + AC] has to be probed quantitatively. Qualitatively, one can attempt to understand it in terms of the behaviour of χ within the framework of the Patterson–Delmas theory [4, 5, 10]. Both water and heavy water constitute a very bad solvent for polystyrene. Its addition to the system can be viewed as acting to lower the solvent quality of acetone. This would imply an increase in the contact energy dissimilarity between the polymer and the solvent and, hence, an

crease in the contribution due to the first term in χ (in eq. (1)) upon the addition of water heavy water to the system. The contribution to the second term in χ (eq. (1)) stems from e free-volume disparity (due to the difference in the degree of thermal expansion) tween the polymer and solvent molecules. The coefficient of thermal expansion (α) of ater and heavy water is much *lower* than that of acetone ($\alpha_W = 0.257 \times 10^{-3}\,\mathrm{deg}^{-1}$, $_{HW} = 0.185 \times 10^{-3}\,\mathrm{deg}^{-1}$ and $\alpha_{AC} = 1.42 \times 10^{-3}\,\mathrm{deg}^{-1}$). For PS, as expected, α has a uch lower value than that of AC, water and heavy water ($\alpha_{PS} = 0.0824 \times 10^{-3}\,\mathrm{deg}^{-1}$). ie addition of water or heavy water should, thus, lead to a reduction in the contribution to originating from the free-volume disparity (second term in eq. (1)). The overall haviour of χ is a consequence of the competition between these two terms (eq. (1)). χ_c pends only on the ratio of the molar volume of the polymer to the molar volume of the lvent (i.e., on the ratio of the two molecular weights). The addition of water or heavy ater does not alter this ratio and, hence, one does not expect χ_c to change drastically. The dition of water/heavy water leads to a net increase in χ (as χ_c remains constant), which, turn, brings about a reduction in ΔT. It can be argued that the desired net *increase* in χ possible only if the *increase* in χ due to the contact energy dissimilarity term far exceeds e *reduction* in χ brought about by the free-volume disparity term. In a gross sense, the dition of water is similar to the reduction in pressure (figure 1) in that it raises χ (leaving constant) and thus reduces ΔT, i.e., the region of miscibility. The fact that X_{CDP} for [PS AC + heavy water] is *slightly* greater than X_{CDP} for [PS + AC + water] (figure 4) may attributed to somewhat lower α for heavy water in comparison to that for water. owever, in the absence of a clear microscopic understanding, the role of water or heavy ater needs to be elucidated further.

An intriguing part of our observations is the absence of one-phase *hole* in *hourglass* nfiguration (figure 6). The nature of the thermodynamic state where this one-phase *le* can be shrunk to a point is anticipated to be tricritical, and it reduces to a triple point the monodisperse limit of the polymer [11]. The behaviour near such a novel ermodynamic state has not yet been experimentally ascertained. However, no one-phase *le* was observed [11] for isotopically substituted [PS + AC]. Perhaps the much lower olecular weight of the polymer in our system could have masked the appearance of this usual feature. Further experimental and analytical work in this direction is required fore any firm conclusions can be reached.

It should be rewarding to compare normal and isotopically substituted [PS + AC + ater] solutions in terms of the dynamics of phase separation, and other contexts [6, 9].

cknowledgements

ie authors are thankful to T Narayanan and E S R Gopal for their advice and couragement. They benefitted immensely from the illuminating and inspiring rrespondence with D Patterson, R L Scott and M Nakata. The cooperation of Jaby cob and T Chacko for style-editing is appreciated. One of the authors (BMJA) thanks e University Grants Commission for the award of a Research Fellowship. The pivotal nancial support for this research by the Department of Atomic Energy and the eparment of Science and Technology is gratefully acknowledged. The authors are debted to the referee for extremely useful suggestions.

References

[1] S Venkatachalam, A Kumar and E S R Gopal, *J. Chem. Phys.* **103**, 6645 (1995)
[2] A G Casielles, R G Rubio, F Monroy and F Ortega, *Phys. Rev.* **E49**, 1404 (1994)
H Hosokawa, M Nakata and T Dobashi, *J. Chem. Phys.* **98**, 10078 (1993)
[3] W Shen, G R Smith, C M Knobler and R L Scott, *J. Phys. Chem.* **94**, 7943 (1990)
T Dobashi and M Nakata, *J. Chem. Phys.* **99**, 1419 (1993)
[4] K S Siow, G Delmas and D Patterson, *Macromolecules* **5**, 29 (1972)
[5] L Zeman and D Patterson, *J. Phys. Chem.* **76**, 1214 (1972)
[6] T Narayanan and A Kumar, *Phys. Rep.* **249**, 135 (1994)
[7] D L Sidebottom and C M Sorensen, *J. Chem. Phys.* **89**, 1608 (1988)
[8] R J Tufeu, P H Keyes and W B Daniels, *Phys. Rev. Lett.* **35**, 1004 (1975)
[9] M Luszczyk, L P Rebelo and W A Van Hook, *Macromolecules* **28**, 745 (1995)
[10] D Patterson, *Macromolecules* **2**, 672 (1969)
[11] L P Rebelo and W A Van Hook, *J. Polym. Sci.* **B31**, 895 (1993)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 5,
November 1997
pp. 505–514

Self-consistent numerical study of pure and impurity doped three-band Hubbard model

PRASENJIT SEN and AVINASH SINGH
Department of Physics, Indian Institute of Technology, Kanpur 208 016, India
Email: prasen@iitk.ernet.in

MS received 24 December 1996

Abstract. The three-band Hubbard model – both pure and with static non-magnetic impurities – has been studied within a self-consistent numerical Hartree–Fock (HF) scheme. The system shows nesting properties only in the absence of direct O–O hopping. Spin excitations in the system are gapless with the existence of a Goldstone mode in the broken-symmetry state. The variation of spin-wave velocity with Cu-site Coulomb repulsion shows a $(1/(2U_d)+(1/\Delta))$ dependence in the strong-coupling limit. Each non-magnetic impurity in the system gives rise to two gap states for a particular spin and the local moment produced is robust even at finite concentration of mobile hole doping. The gapless Goldstone mode is preserved even in case of unequal concentration of impurities on the two sublattices.

Keywords. Three-band Hubbard model; charge-transfer insulator; Mott–Hubbard insulator; high T_c superconductor; impurity.

PACS Nos 71.27; 71.55; 75.10; 75.20

1. Introduction

The three-band Hubbard model has been proposed by many authors [1, 2] as a minimal model for describing the physical properties of the copper-oxide planes in high-temperature superconductors. Extensive studies that have been done on this model were motivated by the observation that the undoped parent compounds of the superconducting cuprates are charge-transfer insulators (CTI) rather than Mott–Hubbard (MH) insulators.

It has been argued that this model can be reduced to an effective one-band Hubbard model in order to describe the low-energy physics of the copper-oxide planes when only the spin degrees of freedom are important. A number of studies of the magnetic properties of the copper-oxide planes have been done within a one-band Hubbard model. But recently, lot of experiments have been done on the high-T_c cuprates doped with static, and in particular non-magnetic impurities like Zn^{2+}, Al^{3+}, Ga^{3+} etc., and although some of the properties of these systems have been understood within a one-band model, others require a more detailed picture like the three-band model for their understanding.

It has been established through susceptibility [3] measurements that non-magnetic impurities substituting Cu in the copper-oxide planes generate local moments. NMR [4, 5] measurements show that moments form on the neighbouring Cu sites. Other effects that these impurities have include a rapid reduction in superconducting transition

temperature (T_c) [4], onset of insulating behavior in the metallic state [6], a reduction in Néel temperature T_N [7] and spin-glass behavior in the insulating state [7]. Microscopic understanding of some of these effects have been gained by studying the effects of static non-magnetic impurities in MH insulators described by one-band Hubbard model. Within a HF description of the Hubbard model, the impurity-problem has been studied both analytically within a T-matrix analysis in the limit of large impurity potential [9] and within a self-consistent numerical procedure on finite-sized systems [10]. In both cases gap-states localized around the impurity sites are produced for a particular spin, depending on the sublattice of impurity substitution. These states, when occupied with electrons, give rise to local moments.

In one-hand Hubbard model in hole picture the gap state is formed slightly below the upper Hubbard band (UHB), between the two bands. This being the lowest unoccupied state for the Cu $3d$ holes, the first added hole goes into this state and the moment is lost. At temperatures comparable to the energy difference between the gap state and the lower edge of the UHB, part of the moment can be recovered due to thermal excitation of the hole from the gap state to the UHB [11]. But how the moment survives at low temperatures and finite concentration of mobile hole doping is not clear.

In this paper we consider the full three-band Hubbard model description of the Cu $3d$ holes in the copper-oxide planes in the insulating state when there is exactly one electron (or equivalently one hole) per copper site and present a self-consistent numerical study of the antiferromagnetic (AF) state. In the absence of direct O–O hopping the pure system shows an antiferromagnetic ordering as soon as there is a non-zero Coulomb repulsion on the Cu sites (U_d). For non-zero O–O hopping a critical value of U_d is required for the AF order to set in. With the generally accepted set of values for the parameters, the system is CTI at the half-filled level. It has three bands—lower and upper Hubbard bands at the bottom and the top and the oxygen band in the middle. Each non-magnetic impurity in the three-band model gives rise to two gap-states—one slightly below the UHB and one below the oxygen band. Oxygen band being lower in energy, the upper gap-state remains occupied even at finite hole dopings making the moment robust.

We have also studied the nature of spin-wave excitation in the system. In the pure case the spin-wave excitation is gapless and one has a Goldstone mode. For $U_d \gg t_{pd}$ the spin-wave velocity increases with decreasing U_d showing roughly a $((1/2U_d) + (1/\Delta))$ dependence as in the equivalent Heisenberg model. The Goldstone mode is preserved even in presence of impurities.

2. Pure system

We consider the following three-band Hubbard Hamiltonian for the $3d$ holes in the copper-oxide planes, which is shown in figure 1 (adapted from ref. [14])

$$H = -t_{pd}^{ij} \sum_{\langle i,j\rangle;\sigma} (p_{i\sigma}^{\dagger} d_{j\sigma} + \text{h.c.}) - t_{pp}^{ij} \sum_{\langle l,m\rangle;\sigma} (p_{l\sigma}^{\dagger} p_{m\sigma} + \text{h.c.}) + \epsilon_p \sum_{i\sigma} p_{i\sigma}^{\dagger} p_{i\sigma}$$
$$+ \epsilon_d \sum_{i\sigma} d_{i\sigma}^{\dagger} d_{i\sigma} + U_d \sum_i n_{i\uparrow}^d n_{i\downarrow}^d + U_p \sum_i n_{i\uparrow}^p n_{i\downarrow}^p + U_{pd} \sum_{\langle i,j\rangle} n_i^p n_i^d, \quad (1)$$

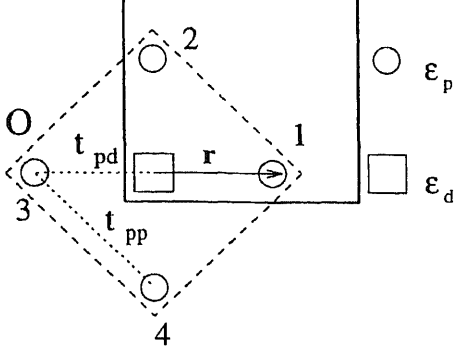


Figure 1. Basis (solid line bounded square) and CuO_4 plaquette (dashed line bounded square) of the CuO_2 plane of the three-band Hubbard model. The squares and the circles mark the Cu $3d_{x^2-y^2}$ and O positions, at energies ϵ_d and ϵ_p, respectively. The $2p_x$ and $2p_y$ orbitals are located at 1,3 and 2,4, t_{pd}^{ij} denotes the transfer integral between copper and oxygen and t_{pp}^{ij} the direct oxygen hopping. The lattice spacing $a = 2|r|$.

where $p(p^\dagger)$ and $d(d^\dagger)$ are the annihilation (creation) operators for holes on the O and Cu sites, $\langle i, j\rangle$ are nearest neighbour (nn) Cu and O sites and $\langle l, m\rangle$ are two neighboring O sites; n's are the respective number densities. t_{pd} is the hopping term between neighboring Cu $3d$ and O $2p$ orbitals and t_{pp} is the hopping matrix element between two nn O $2p$ orbitals. ϵ_d and ϵ_p are the on-site energies of a hole sitting on Cu $3d$ or O $2p$ orbital. U_d and U_p are the Coulomb repulsion terms on the Cu and O sites respectively. U_{pd} is the repulsion between charges on the neighbouring Cu and O sites. σ is the spin index. The so called charge-transfer gap Δ denotes the energy difference between oxygen and copper, $\Delta \equiv \epsilon_p - \epsilon_d$. In the hole picture Δ is positive. The orbital symmetries of the type $d_{x^2-y^2}$ and $p_{x,y}$ imply additional phase factors for the hopping integrals t_{pd}^{ij} and t_{pp}^{ij},

$$t_{pd}^{0j} = \phi^j t_{pd}, \quad \phi^1 = \phi^2 = -1, \quad \phi^3 = \phi^4 = 1,$$
$$t_{pp}^{ij} = \psi^{ij} t_{pp}, \quad \psi^{12} = \psi^{34} = -1, \quad \psi^{23} = \psi^{41} = 1,$$

where 0 labels the central copper and 1, 2, 3, 4 the four oxygen sites on the plaquette. In this section we shall be talking about a half-filled system when we have exactly one hole per copper site.

For all our calculations we take $t_{pd} = 1.3\,\text{eV}$, $\Delta = 3.0\,\text{eV}$, $U_p = 0\,\text{eV}$ and $U_{pd} = 0\,\text{eV}$ and a 10×10 CuO_2 lattice. A more detailed discussion on the choice of the values of parameters is given in the section on doped system. In the HF approximation relevant to symmetry breaking along z-direction only, the interaction terms can be written as $U\langle n_{i\uparrow}\rangle n_{i\downarrow} + U\langle n_{i\downarrow}\rangle n_{i\uparrow}$, so that the on-site energy of spin $\bar{\sigma} \equiv -\sigma$ hole on the ith Cu site is $U\langle n_{i\sigma}\rangle$. In the site basis, matrix elements of the Hamiltonian for spin σ are given by $\langle i \mid H_\sigma \mid i\rangle = \epsilon_d + U_d\langle n_{i\bar{\sigma}}\rangle$ if i is a Cu site and $\langle i \mid H_\sigma \mid i\rangle = \epsilon_p + U_p\langle n_{i\bar{\sigma}}\rangle$ if i is an oxygen site. $\langle i \mid H_\sigma \mid j\rangle = \pm t_{pd}$ if i and j are nn Cu and O sites and $\langle l \mid H_\sigma \mid m\rangle = \pm t_{pp}$ if l and m

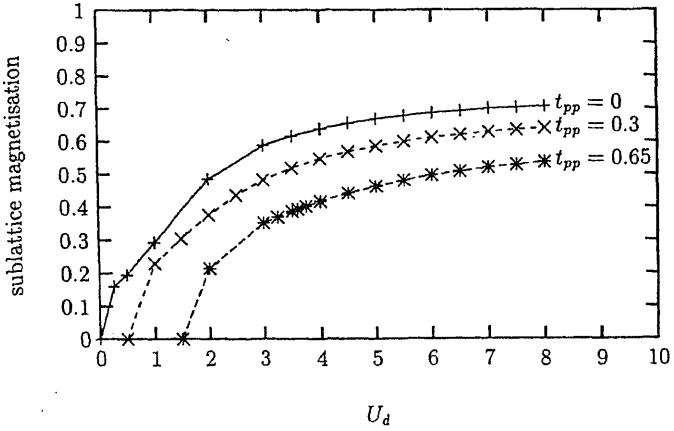


Figure 2. Sublattice magnetization as a function of U_d for different values of t_{pp} (all energies in eV).

are two nn O sites. Choice of an initial configuration of spin $\bar{\sigma}$ densities on the Cu sites gives the starting Hamiltonian matrix H_σ which is diagonalized to give the eigensolutions $E_{l\sigma}$, $\phi_{l\sigma}$. From these the spin densities are evaluated using $\langle n_{i\sigma} \rangle = \Sigma_{E_{l\sigma} < E_F} (\phi^i_{l\sigma})^2$. H_σ is updated from these computed spin densities and the procedure is continued till self-consistency is achieved.

Exact numerical (QMC) calculations reveal that the three-band Hubbard model with a finite positive U_d possesses an AF ground-state [8]. In order to study the magnetic behavior of the system with changing U_d we have calculated the sublattice magnetization as a function of U_d. Figure 2 shows the variation of sublattice magnetization as a function of U_d for different values of t_{pp}. For $t_{pp} = 0$ the system goes into an AF state for any finite positive U_d. This is because of nesting of the Fermi surface. When $t_{pp} \neq 0$ the system loses its nesting property and a critical value of U_d is required for the system to become antiferromagnetic. This critical value of U_d increases with t_{pp}.

A study of the energy spectrum of the system shows that for large values of U_d ($U_d \gg \Delta$) the system is a CTI. For an $L \times L$ CuO_2 lattice, having L^2 Cu and $2L^2$ O sites, the first $L^2/2$ states form the lower Hubbard band (LHB). The oxygen $2p$ band in the middle has $2L^2$ states and the UHB has another $L^2/2$ states. The $2p$ band is further split into three sub-bands similar to Zhang–Rice singlet-triplet splitting. The singlet sub-band lies below the bare oxygen energy, there are a group of non-bonding (NB) states at the bare oxygen energy ϵ_p and the triplet sub-band lies above these NB states. For $U_d < \Delta$ but above the critical value so that the system is an AF, the UHB goes below the oxygen band. The charge-gap is maintained and the system is a MH insulator. Below the critical value of U_d when the sublattice magnetization vanishes, the charge-gap also disappears and the system becomes a d-type metal. There are no singlet-triplet splitting in the oxygen band in these cases (figure 3). The situation in intermediate range of values of U_d is more complicated because the UHB and the oxygen band are strongly mixed.

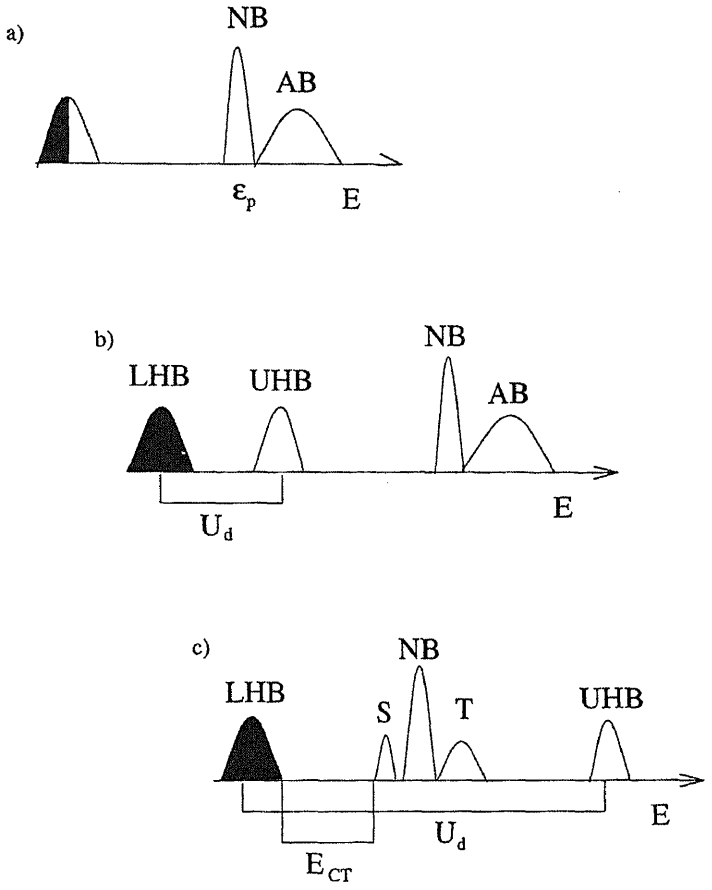


Figure 3. Single-particle spectra for the three-band Hubbard model in the hole picture: **(a)** metal, **(b)** Mott–Hubbard insulator, **(c)** charge-transfer insulator with a singlet-triplet splitting. Shaded regions denote the occupied states. $[N](A)B$ = [non] (anti)bonding, S = singlet, T = triplet, E_{CT} = renormalised charge-transfer gap.

Spin waves are the low-energy collective excitations representing the transverse spin fluctuation in the AF ground state. In the one-band Hubbard model and its strong-coupling analogue, the Heisenberg model, the nature of spin excitations has been studied extensively. It is well known that in these systems the spin-wave spectrum is gapless with the existence of a Goldstone mode. The dispersion for the long-wavelength spin-wave modes is linear in the momentum Q. The spin-wave modes are obtained from the poles of the time-ordered transverse spin propagator $\langle T[S^-(i,t)S^+(j,t')]\rangle$ which, in the random phase approximation (RPA), has the form $[\chi^{-+}(\Omega)] = [\chi_0^{-+}(\Omega)]/(1 - U[\chi_0^{-+}(\Omega)])$. The matrix elements of $[\chi_0^{-+}(\Omega)]$, the bare antiparallel-spin particle-hole propagator, are given by

$$[\chi_0^{-+}(\Omega)]_{ij} = i\int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, g_{ij}^{\uparrow}(\omega) g_{ji}^{\downarrow}(\omega - \Omega), \tag{2}$$

where g's are the Green's functions in the AF state. In terms of the eigensolutions E_l, ϕ_l in the self-consistent state $\chi_0^{-+}(\Omega)$ has the form

$$[\chi_0^{-+}(\Omega)]_{ij} = \sum_{E_{l\uparrow}>E_F}^{E_{m\downarrow}<E_F} \frac{\phi^i_{l\uparrow}\phi^j_{l\uparrow}\phi^i_{m\downarrow}\phi^j_{m\downarrow}}{E_{l\uparrow} - E_{m\downarrow} - \Omega} + \sum_{E_{l\uparrow}<E_F}^{E_{m\downarrow}>E_F} \frac{\phi^i_{l\uparrow}\phi^j_{l\uparrow}\phi^i_{m\downarrow}\phi^j_{m\downarrow}}{E_{m\downarrow} - E_{l\uparrow} + \Omega}. \tag{3}$$

Since only the Cu spins form AF order, spin-waves will have amplitude only on the Cu sites, and so we calculate $[\chi_0^{-+}(\Omega)]_{ij}$ connecting only the Cu sites. In (3) the site indices i and j thus run only through the Cu sites, but the eigenvalue indices l and m run through all three bands.

All information regarding the nature of the spin-wave excitations is contained in the eigensolutions of $[\chi_0^{-+}(\Omega)]$, as is evident from the expansion of the full RPA susceptibility given below in terms of the eigenvalues $\lambda(\Omega)$ and the eigenvectors $|\phi_\lambda(\Omega)\rangle$ of $[\chi_0^{-+}(\Omega)]$. The spin-wave energies are obtained from the pole $1 - U\lambda(\Omega) = 0$ in the RPA spin susceptibility, and the eigenvector yields the spin-wave amplitude

$$[\chi^{-+}(\Omega)] = \sum_\lambda \frac{\lambda|\phi_\lambda(\Omega)\rangle\langle\phi_\lambda(\Omega)|}{1 - U\lambda(\Omega)}. \tag{4}$$

In the three-band model, as in the one-band model, the largest eigenvalue $\lambda_{\max}(\Omega = 0)$ of the $[\chi_0^{-+}(\Omega)]$ matrix for $\Omega = 0$ is found to be exactly equal to $1/U_d$, confirming that the spin-wave excitations are gapless.

The higher (nth) spin-wave mode energies Ω_n are determined by solving $1 - U\lambda_n(\Omega_n) = 0$ for the appropriate (nth from the top) eigenvalue $\lambda_n(\Omega)$. The root of the equation $1 - U\lambda_n(\Omega_n) = 0$ is determined by obtaining $\lambda_n(\Omega)$ for several close values of Ω on both sides of the root and then linearly interpolating between them. Thus if $(\lambda_n^1, \Omega_n^1)$ and $(\lambda_n^2, \Omega_n^2)$ are two sets of values of two Ω's very close to and on either side of the root, then the root Ω_n is determined from

$$\lambda(\Omega_n) = \frac{1}{U} = \lambda_n^1 + \frac{\lambda_n^2 - \lambda_n^1}{\Omega_n^2 - \Omega_n^1}(\Omega_n - \Omega_1). \tag{5}$$

The procedure involves first obtaining the self-consistent state for the system and then the process of constructing and diagonalizing the $[\chi_0^{-+}(\Omega)]$ matrix for different values of Ω. From the set of eigenvalues $\lambda_n(\Omega)$, the spin-wave energies are then obtained as above. Nature of spin-wave modes can be studied by obtaining the angles of rotation of local spin vectors from the spin-wave amplitudes and the local magnetization S_z^i using [12]

$$\theta_i = \sin^{-1}\frac{\phi^i_\lambda}{S^i_z}. \tag{6}$$

The $\Omega_0 = 0$ mode corresponds to a spin-wave excitation which has equal amplitudes on all sites and rotates all the spins by equal angles and is identified as the Goldstone mode. Thus, as in the one-band model, the spin-wave excitation in a three-band model is gapless and there exists a Goldstone mode as one would expect in a system with continuous spin-rotational symmetry.

Spin-wave modes can be assigned mode numbers [10] which yield wave vectors $Q_{x/y} = (2\pi/L)n_{x/y}$. The spin-wave energy is found to be proportional to the wave vector

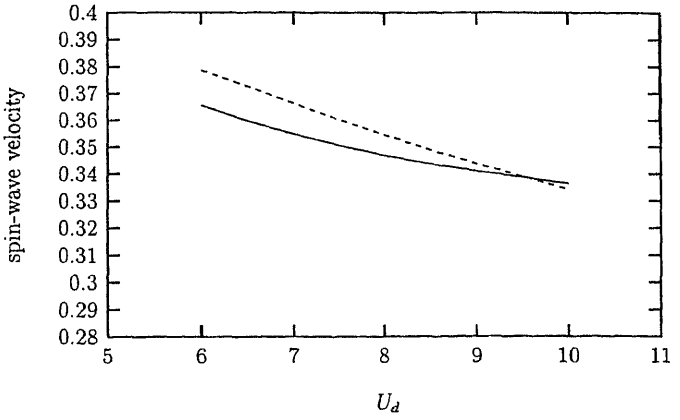


Figure 4. Spin-wave velocity (in eV) obtained at the RPA level for the finite-size system (dashed line) and calculated from analytical results of literature (continuous line) as a function of U_d (in eV).

for the long-wavelength modes (modes for a few λ's at the top) in the strong-coupling ($U_d \gg t_{pd}$) limit. The spin-wave velocity v can be defined by $\Omega_Q = vQ$ in the long-wavelength limit, where $Q = (2\pi/L)\sqrt{n_x^2 + n_y^2}$ is the magnitude of the wave vector. The spin-wave energy scale in the problem is set by the exchange coupling J between nn Cu atoms in the equivalent Heisenberg model to which the three-band Hubbard model can be mapped in the limit $U_d \gg t_{pd}$ and $\Delta \gg t_{pd}$. The spin-wave velocity is then equal to $\sqrt{2}Ja$, where a is the lattice constant, set equal to unity henceforth. The effective exchange coupling for a three-band Hubbard model is given by [13]

$$J = \left(\frac{2t_{pd}^4}{\Delta^2}\right)\left(\frac{1}{2U_d} + \frac{1}{\Delta}\right). \tag{7}$$

In figure 4 we plot the spin-wave velocity, calculated from the energy of the first spin-wave mode, for which $n_x = 1, n_y = 0$ (or equivalently $n_x = 0, n_y = 1$), and also $\sqrt{2}J$ calculated from eq. (7). For large values of U_d the spin-wave velocity calculated from our model is found to be in reasonable agreement with the analytical result. The slight disagreement is not surprising as the condition $\Delta \gg t_{pd}$ is not really satisfied in this case. On a finite lattice the Q-space is discrete and the first mode corresponds to a Q of 0.62 for a 10×10 system. Below U_d of 6.0, spin-wave energy no longer remains linear for this Q and calculation of spin-wave velocity from the simple relation $\Omega_Q = vQ$ is not possible.

3. Impurity-doped system

We have studied the problem of static non-magnetic impurities doped in the three-band Hubbard model. As in the one-band case we model the impurities by putting a high repulsive potential on the impurity (Cu) site [9] to restrict the Cu $3d$ holes from

occupying that site. The Hamiltonian now is

$$H = H_{\text{pure}} + V \sum_{I,\sigma} a^{\dagger}_{I\sigma} a_{I\sigma}, \tag{8}$$

where H_{pure} is the Hamiltonian for the pure system given by (1) and I runs over the impurity sites.

The values of most of the parameters of the model are related to electron spectroscopy and are obtained from comparison with the results of these experiments. Different kinds of theoretical techniques give slightly different sets of values. But a generally acceptable set of values is [14] $t_{pd} = 1.3\,\text{eV}$, $t_{pp} = 0.65\,\text{eV}$, $U_d = 8.8\,\text{eV}$, $U_p = 0\,\text{eV}$, $U_{pd} = 0\,\text{eV}$, $\Delta = 3.0\,\text{eV}$. In this section we work with this set of values of the parameters.

Starting from the pure system at half-filling one Cu $3d$ hole is removed for each added impurity. For each impurity, depending on the sublattice of its substitution, a self-consistent solution of this Hamiltonian shows that in the spectrum of one particular spin there are three states outside the bands and for the other there is one outside. Consider an impurity on an A-sublattice site (having majority of $\uparrow$-spin holes), in an $L \times L$ system. In the $\uparrow$-spin spectrum there are $(L^2/2 - 1)$ states in the LHB. One state goes out to the defect state which is the highest in energy ($\sim V$). The oxygen band and the UHB have $(2L^2 - 1)$ and $(L^2/2 - 1)$ states respectively, one state from each going out to form the gap states. The $L^2/2$th and the $2L^2$th states are the two gap states. In the spectrum for the $\downarrow$-spin hole, the LHB and the oxygen band remain unaffected. The UHB has $(L^2/2 - 1)$ states, one state going to the defect state. The situation is simply reversed if the impurity is placed on a B-sublattice site. Coming to the nature of these states, we find that the defect state is essentially site-localized on the impurity site, whereas the lower gap state, coming out of the oxygen band, has amplitude mainly on the oxygen sites around the impurity, and the upper gap state has amplitude mainly on the four Cu sites neighbouring the impurity. In this respect the upper gap state is similar to the gap state in a one-band model [9, 10]. This state, when occupied by electrons, creates a local moment residing predominantly on the four Cu sites neighbouring the impurity. The defect state has amplitude primarily on the impurity site. With only a non-magnetic impurity on an A-sublattice and no extra added hole, the system has $(L^2/2 - 1)\uparrow$-spin and $L^2/2\downarrow$-spin holes. So the LHB is completely filled and all the other states are empty of holes. The lower gap state being the lowest unoccupied state for the holes, the first added $\uparrow$-spin hole goes into this state and subsequent holes go into the oxygen band. So even at finite concentration of mobile hole doping the upper gap state remains occupied by electrons. Thus unlike in one-band model, the local moment in three-band model is robust with respect to mobile hole doping. In figure 5 we show the variations of the two energy gaps—between lower gap-state and oxygen band and the upper gap-state and the UHB. It is interesting to note that although the lower gap varies considerably, the upper gap is relatively insensitive to the variation of V.

The magnitude of the local moment is proportional to the density of the gap state on the copper sites. We find that the total density of the upper gap state on the copper sites is less than 1. Part of the density goes out to the oxygen sites because of the Cu–O hybridization. This explains the observation that the measured moment is less than the theoretically expected value [3]. If one spin is removed from a compensated spin system like an AF, the resulting moment should be that due to a spin-$\frac{1}{2}$ electron, which is

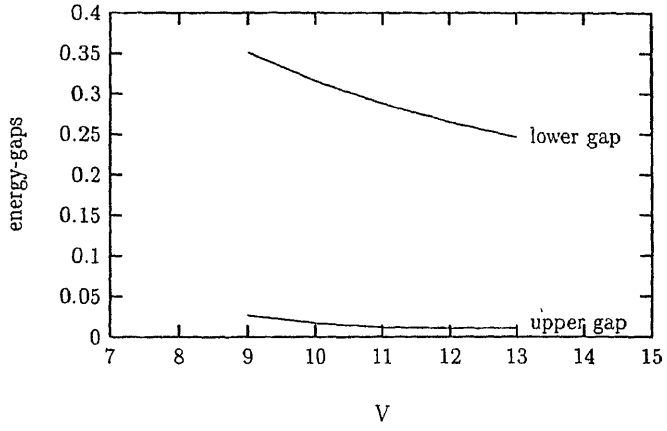


Figure 5. Variation of the two energy-gaps with V (all in eV).

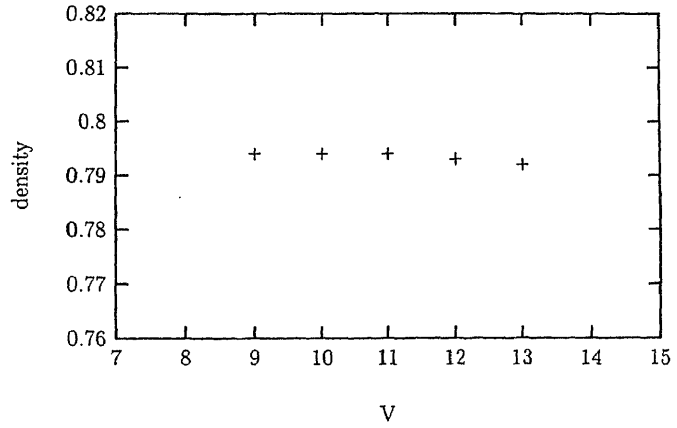


Figure 6. Gap-state density on the Cu-sites for different impurity potential V (in eV).

$(g\sqrt{s(s+1)} =)1.9\,\mu_B$. But in experiments the moment is measured to be $1.2\,\mu_B$, about 35% less than the calculated value. More importantly this is found to be independent of the nature of the non-magnetic impurity, Zn^{2+}, Al^{3+}, Ga^{3+}—all are seen to give rise to moment of about the *same* magnitude. In our calculations we find that about 21% of the gap-state density escapes to the oxygen sites from copper. Figure 6 shows total density of the second gap state on Cu sites which is interestingly found to be essentially constant over the range of V studied in our calculations. This leads to the significant conclusion that the gap-state-induced moment is more or less independent of the details of the non-magnetic impurity within our model of the impurity potential term.

We have also studied the nature of spin-wave excitations in the impurity-doped system. We find that the gapless Goldstone mode associated with spontaneous symmetry-breaking in the AF ground state is preserved in presence of impurities and even when the concentration of impurities on the two sublattices are not equal. This is because the $SU(2)$

symmetry of the Hamiltonian, which is broken in the AF ground state, is still preserved. Wan *et al* [15] have obtained two modes of spin-wave excitation in case of unequal concentration of impurities on the two sublattices. We think the gapless mode in our calculation actually corresponds to their acoustic mode.

4. Conclusion

We have studied the pure and impurity-doped three-band Hubbard models. When direct O–O hopping is not present the pure system shows nesting property and goes into an AF state for any non-zero U_d. In presence of a finite t_{pp} the nesting is lost and the system becomes an AF only above a certain critical value of U_d, which increases with t_{pp}. The transverse spin excitations in the system have a gapless Goldstone mode. The spin-wave velocity, in the strong-coupling limit, shows a $((1/2U_d) + (1/\Delta))$ dependence in agreement with the strong-coupling perturbative approach.

In the doped system each non-magnetic impurity produces two gap states. The upper gap state remains occupied by electrons even in presence of finite concentration of mobile holes so that the local moment formed is robust. The density of the occupied gap state on the Cu sites is less than 1, so that the local moment is less than the theoretical value of $1.9\,\mu_B$. The local moment is found to be insensitive to the impurity-potential V suggesting that the observed moment is independent of the detailed nature of the impurity. Even with unequal concentration of impurities on two sublattices, the spin-wave excitation remains gapless, and the Goldstone mode is also preserved.

References

[1] V J Emery, *Phys. Rev. Lett.* **58**, 2794 (1987)
[2] C M Varma, S Schmitt-Rink and E Abrahams. *Solid State Commun.* **62**, 681 (1987)
[3] G Xiao, M Z Cieplak, J Q Xiao and C L Chien, *Phys. Rev.* **B42**, 8752 (1990)
[4] R E Walstedt, R F Bell, L F Schneemeyer and J V Waszczak. *Phys. Rev.* **B48**, 10646 (1993)
[5] A V Mahajan, H Alloul, G Collin and J F Marucco, *Phys. Rev. Lett.* **72**, 3100 (1994)
[6] G Xiao, M Z Cieplak, A Gravin, F H Stritz, A Bakhshai and C L Chien, *Phys. Rev. Lett.* **60**, 1446 (1988)
[7] B Keimer, A Aharony, A Auerbach, R J Birgeneau, A Cassanho, Y Endoh, R W Erwin, M A Kastner and G Shirane, *Phys. Rev.* **B45**, 7430 (1992)
[8] G Dopf, A Maramatsu and W Hanke, *Europhys. Lett.* **17**, 559 (1992)
[9] P Sen, S Basu and A Singh, *Phys. Rev.* **B50**, 10381 (1994)
[10] S Basu and A Singh, *Phys. Rev.* **B53**, 6406 (1996)
[11] P Sen and A Singh (unpublished)
[12] A Singh and Z Tešanović, *Phys. Rev.* **B41**, 614 (1990)
[13] J Zaanen and A Olés, *Phys. Rev.* **B37**, 9423 (1988)
[14] W Brenig, *Phys. Rep.* **251**, 179 (1995)
[15] C C Wan, A B Harris and D Kumar, *Phys. Rev.* **B48**, 1036 (1993)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 5,
November 1997
pp. 515–519

Measurement of fast neutron induced fission cross section of thorium using Lexan plastic track detector

R K JAIN, H S VIRK, J RAMA RAO* and S K BOSE*
Physics Department, Guru Nanak Dev University, Amritsar 143 005, India
*Physics Department, Nuclear Research Laboratory, Banaras Hindu University, Varanasi 221 005, India

MS received 14 February 1997; revised 3 July 1997

Abstract. Fission-track registration characteristics of Lexan solid state nuclear track detectors have been used to measure the fast neutron induced fission cross section of ^{232}Th. The fast neutrons ($\cong$ 14.2 MeV) were produced with the help of an AN-400 model Van-de-Graaff accelerator at Banaras Hindu University laboratory using $^{3}H(^{2}H, n)^{4}He$ reaction and were used to irradiate the fissile target deposited on the plastic detector. The track density T, registered on the plastic detector is related to the fission cross section σ_f, through the relation $T = kn\sigma_f\phi t$ where n is the number of fissile atoms per cm^2 in the deposit, ϕ is the neutron flux, k is fission track registration efficiency and t is the time of irradiation. The fission cross section σ_f of ^{232}Th, relative to the well measured fission cross section of ^{238}U, was found to be 0.36 ± 0.04 barn.

Keywords. Nuclear track detectors; induced fission; cross section; neutron flux; track density; Lexan; thorium.

PACS No. 25.85

1. Introduction

Walker *et al* [1] were the first to suggest that the fluence of neutrons can be measured by counting the fission tracks in a thin, track-recording sheet in contact with a foil of fissionable material. This method has a number of attractions. The foremost of these is that single fission produces a large amount of energy ($\sim$ 200 MeV) which is not very dependent on the energy of the neutron producing the reaction [2, 3]. Most of this energy ($\sim$ 160 MeV) is shared between two easily detected fission fragments which travel in opposite directions and which for most nuclides and neutrons ($E_n \sim 7$ MeV), are approximately isotropically distributed with respect to the direction of the incident neutron [4, 5]. Fission fragment 'tracks' can be seen with an optical microscope in a number of insulating materials (SSNTDs) when properly etched [6]. By choosing a suitable material, for example, Lexan plastic, the fission events can be recorded with essentially a zero background from competing nuclear events.

Taking advantage of these attractive features, in the present work, the fission cross section of ^{232}Th has been determined relative to the standard fission cross section of ^{238}U by comparing the fission track densities produced in Lexan plastic in both the cases. The

method employed uses a thin deposit of thorium solution on the Lexan plastic in a 2π geometry [7].

2. Experimental details

Several pieces (2 cm × 2cm) of Lexan plastic were cut from a sheet of uniform thickness ($\sim 200\,\mu m$). A known weight (1 gm of $UO_2(NO_3)_2 \cdot 6H_2O$) of the fissile material uranyl nitrate was first dissolved in $10\,cm^3$ of distilled water and a known amount ($0.25\,cm^3$) of this solution was deposited on five Lexan plastic foils and dried with the help of an infrared lamp. In the same way thorium nitrate ($ThO_2(NO_3)_2 \cdot 6H_2O$) solution under the same condition was prepared and the same amount ($0.25\,cm^3$) of the solution was deposited on five other Lexan plastic foils. It was ensured that the thickness of fissile material in each case was so small that the errors due to self-absorption effects are negligible. These coated plastics were exposed to the fast neutrons for different lengths of time ranging from one hour to several hours. The corresponding neutron fluences varied from 1×10^{10} to $6.7 \times 10^{10}\,n/cm^2$. All the pieces of the Lexan plastics were irradiated with 14.2 MeV neutrons produced with the help of a 400 kV Van de Graaff accelerator, at a distance of about 1.5 cm from the place of the tritium target. Tritium target was at right angle to the deutron beam. Both the uranium and thorium targets are irradiated together using the same fluence. For an independent determination of the flux, an aluminium foil of the same area as that of target was also irradiated sandwiched between the thorium and uranium targets. After irradiation, the plastic pieces along with an unirradiated (control) plastic were etched in 6.25 N NaOH solution at 60°C for 1–2 hours in a thermostatically controlled oven. Etched fission fragment tracks in plastics were counted to determine track density by viewing under an optical microscope (Olympus, Model BH-2) using a magnification of 600×. In each plastic foil, equal area were scanned ($6.25 \times 10^{-2}\,cm^2$). During scanning, proper care was taken not to count the same area of the detector more than once in any case.

3. Results and discussion

The neutron flux, as monitored by the gamma spectrometric measurement of irradiated ^{27}Al foil at the site where all the detector pieces were irradiated was found to be typically $3 \times 10^6\,n\,cm^{-2}s^{-1}$. Table 1 shows the irradiation time and number of tracks in $6.25 \times 10^{-2}\,cm^2$ for uranium and thorium and table 2 shows the irradiation time, fluence and track density for uranium and thorium. The microphotographs of fission tracks registered in Lexan plastic for uranium and thorium are shown in figures 1 and 2 respectively. Figure 3 shows the fission track density vs. neutron fluence for uranium and thorium. It can be seen that the track density varies linearly as the fluence and the slope of the straight line gives track density per unit fluence. Thus the ratio of the slopes of the two lines, δ_U/δ_{Th}, gives the ratio of track densities T_U/T_{Th}, on a relative basis. Since the detector is in contact with the fissile material in 2π geometry one has for ^{238}U and ^{232}Th,

$$T_U = (k\phi t) n_U \sigma_{fU}, \tag{1}$$

$$T_{Th} = (k\phi t) n_{Th} \sigma_{fTh}. \tag{2}$$

Table 1. Number of fission fragments measured in Lexan plastics for uranium and thorium.

Irradiation time (hrs)	Number of tracks for ^{238}U	Number of tracks for ^{232}Th
1.00	1127	375
2.25	2556	856
3.00	3412	1137
4.50	5162	1700
6.50	7387	2462

Scanned area is $6.25 \times 10^{-2}cm^2$.

Table 2. Track density measured in Lexan plastic for both uranium and thorium.

Irradiation time (hrs)	Fluence 10^{10} n/cm^2	Track density for $^{238}U(\times 10^4/cm^2)$	Track density for $^{232}Th(\times 10^4/cm^2)$
1.00	1.0	1.82	0.60
2.25	2.3	4.09	1.37
3.00	3.1	5.46	1.82
4.50	4.6	8.17	2.72
6.50	6.7	11.82	3.94

Figure 1. Microphotograph of fission tracks found in Lexan plastic coated with uranium, irradiated with fast neutrons for 6.5 hrs.

Here, the factor $k\phi t$, the product of efficiency, flux and irradiation time, is the same for both cases, because they were irradiated together under the same conditions. In (1) σ_{fU} is fission cross section for ^{238}U and n_U is number of atoms per cm^2 of ^{238}U and in (2) σ_{fTh} is fission cross section for ^{232}Th and n_{Th} is number of atoms per cm^2 of ^{232}Th. From eqs (1) and (2)

$$\frac{T_U}{T_{Th}} = \frac{n_U \sigma_{fU}}{n_{Th} \sigma_{fTh}}. \tag{3}$$

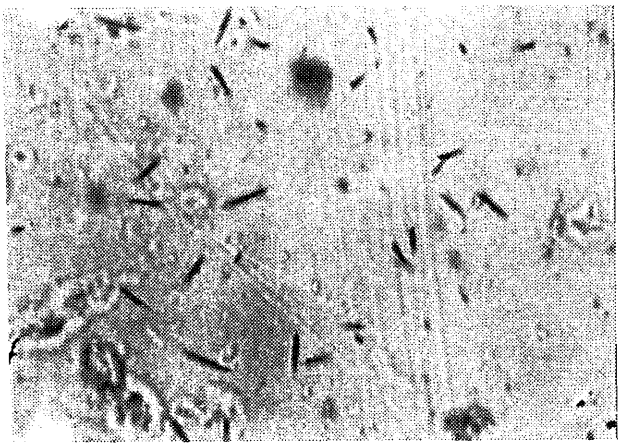


Figure 2. Microphotograph of fission tracks found in Lexan plastic coated with thorium, irradiated with fast neutrons for 6.5 hrs.

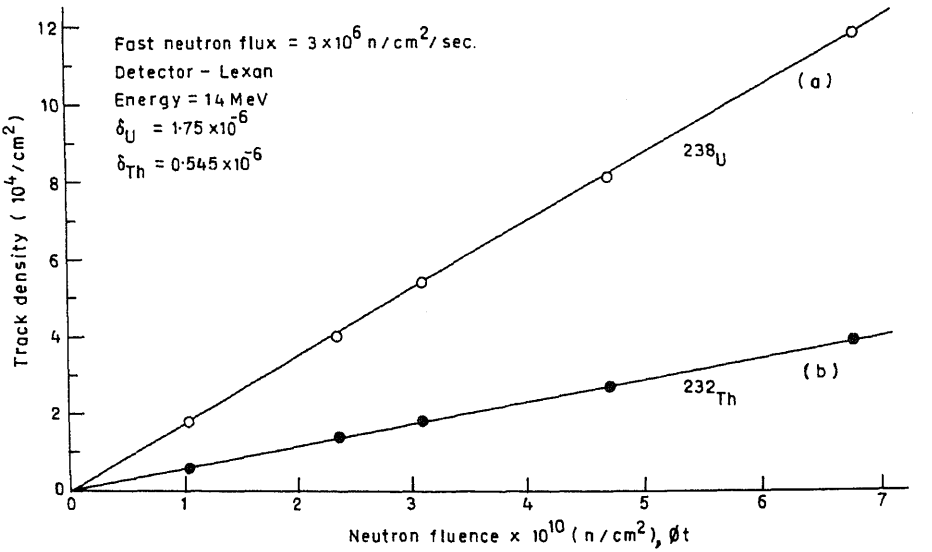


Figure 3. Variation of fission track density vs. neutron fluence for Lexan coated with (a) ^{238}U and (b) ^{232}Th.

The value of δ_U/δ_{Th} was obtained from figure 3 as explained above. The ratio, n_U/n_{Th}, was determined from the relative weights of the uranium and thorium deposits both having the same area. n_T (number of atoms per cm^2 of the fissile target) can be defined by

$$n_T = \frac{W_i N_a P_i}{A_i}, \tag{4}$$

where W_i is the weight of the fissile material per unit area, P_i is the isotopic abundance of the target isotope and N_a is the Avagadro's number (6.03×10^{23} per gm mole). Thus the

fission cross section of ^{232}Th is given by

$$\sigma_{f\,Th} = \frac{n_U \sigma_{f\,U} \delta_{Th}}{n_{Th} \delta_U}. \quad (5)$$

In this way the fission cross section of ^{232}Th was found to be 0.36 ± 0.04 barn, with the standard fission cross section of ^{238}U taken as 1.205 ± 0.020 barn from White and Warner [8]. The present experimental value of fission cross section of thorium by 14.2 MeV neutrons, agrees well within errors with earlier value of 0.34 ± 0.03 barn, obtained by Chowdhuri *et al* [9], using Lexan detector but by a different method of studying the angular distributions and integrating them.

References

[1] R M Walker, P B Price and R L Fleischer, *Appl. Phys. Lett.* **3**, 28 (1963)
[2] R D Evans, *The Atomic Nucleus* (McGraw Hill, New York, 1955) ch. 11
[3] H A Enge, *Introduction to nuclear physics* (Addison Wesley, New York, 1966) ch. 14
[4] R L Henkel, *Fission by fast neutrons in 'Fast Neutron Physics'* edited by J B Marion and J L Fowler (Part 2, John Wiley, Chichester, 1963) pp. 2001–2050
[5] J E Simmons and R L Henkel, *Phys. Rev.* **120**, 198 (1960)
[6] R L Fleischer, P B Price and R M Walker, *Ann. Rev. Nucl. Sci.* **15**, 1 (1965a)
[7] R H Iyer, R Sampath Kumar and N K Chaudhuri, *Nucl. Instrum. Methods* **115**, 23 (1974)
[8] P E White and G P Warner, *J. Nucl. Energy* **21**, 671 (1967)
[9] N K Chaudhuri, V Natrajan, R Sampath Kumar, M L Sagu and R H Iyer, *Nucl. Tracks* **3**, 69 (1979)

AMANA
ournal of
physics

© Printed in India

Vol. 49, No. 5,
November 1997
pp. 521–533

CORPION: A system for coincidences between recoil id projectile ions at NSC, New Delhi

J SINGH, S K GOEL, R SHANKER, D O KATARIA*, N MADHAVAN*, SUGATHAN*, J J DAS*, D K AVASTHI* and A K SINHA*
omic Physics Laboratory, Department of Physics, Banaras Hindu University, Varanasi 221 005, ia
uclear Science Centre, P.O. Box 10502, New Delhi 110 067, India
ail: rshanker@banaras.ernet.in

received 7 December 1996; revised 1 July 1997

stract. An on-line facility to measure coincidences between the recoil ions and the scattered jectiles (SCORPION) has been designed, fabricated and commissioned at Nuclear Science ntre (NSC), New Delhi. The facility consists of a four jaw slit assembly, a time of flight (TOF) ctrometer, a parallel plate electrostatic charge analyser and a one dimensional position sensitive allel plate avalanche counter (PPAC). Details of the design and working principles of various mponents and the test results obtained for the Si^{q+}–Ar collision system are presented to highlight performance of the system. A multiple loss of up to four electrons has been observed for MeV Si^{4+} ions colliding with argon atoms in a single collision condition. Spectra of recoil ions ected in coincidence with a particular charge state of the scattered projectile show a bell shaped ribution as a function of the recoil charge state (r) for the electron loss events. However, the d of recoil ions drops as r increases for the direct ionization channel. Also for electron loss, the k of the recoil ion distribution is seen to shift to a higher recoil charge state as the number of lost ctrons from the projectile increases.

ywords. Coincidence technique; recoil ions; time of flight spectrometer.

CS Nos **07.75; 34.50**

Introduction

incidence measurements between the scattered projectile ions and the recoil ions have racted the attention of many workers [1–4] since the inception of the field by Cocke . These measurements help in understanding the multielectron processes which occur ring the collisions of the type

$$A^{q+} + B \rightarrow A^{(q-i)+} + B^{p+} + (p-i)e^{-}. \quad (1)$$

ch studies not only provide an indepth information on the various fundamental ocesses (direct ionization, e^{-}-capture and e^{-}-loss) responsible for the production of ltiply ionized recoil ions but also help in the understanding of the dynamics of such lisions. The scattered projectile ions $[A^{(q-i)+}]$ in the exit channel of the process scribed by eq. (1) have a mixture of charge states because of the finite probability of

the occurrence of all the fundamental processes mentioned above. For the direct ionization case, $i = 0$, i.e. the charge state of the scattered projectile is identical to that of the incident one. For the electron capture case, $i > 0$ implying thereby that the charge state of the scattered projectile is less as compared to the charge state of the incident projectile depending on the number of electrons it has captured from the target atom during collision. For the electron loss case, $i < 0$, incident-ion loses additional electrons. Recently, a similar attempt in this direction has been pursued at the 15 UD Pelletron accelerator facility of Nuclear Science Centre (NSC) [6], New Delhi. Some measurements using a time-of-flight spectrometer (TOF) have been performed earlier, the details of which are presented elsewhere [7]. In these measurements the selection of the final charge state of the scattered projectile was not possible. As mentioned above, the separation of the various charge states of the scattered projectile is important to identify the channels leading to the recoil ion distributions for the various ionization processes. This can be achieved either by magnetic deflection [2, 4, 8] or by using a parallel plate electrostatic charge analyser [1, 3, 9]. A new facility, SCORPION (System for COincidences between the Recoil and Projectile IONs), which uses a parallel plate electrostatic charge analyser, has been developed and set up at the zero degree exit port of the general purpose scattering chamber (GPSC) [10] beam line of NSC. The set up has been used to perform coincidence measurements between the recoil ions and the charge selected scattered projectiles. This paper reports on the various hardware components of the SCORPION and their design details. A few typical results obtained from the data on Si^{q+}–Ar collisions using the SCORPION facility are presented and discussed.

2. Hardware details of SCORPION

The SCORPION set up for performing experiments on recoil ion projectile ion coincidences is shown in figure 1. The charge state of the recoil ions produced as a result of the interaction in a crossed beam geometry between the highly charged energetic ions

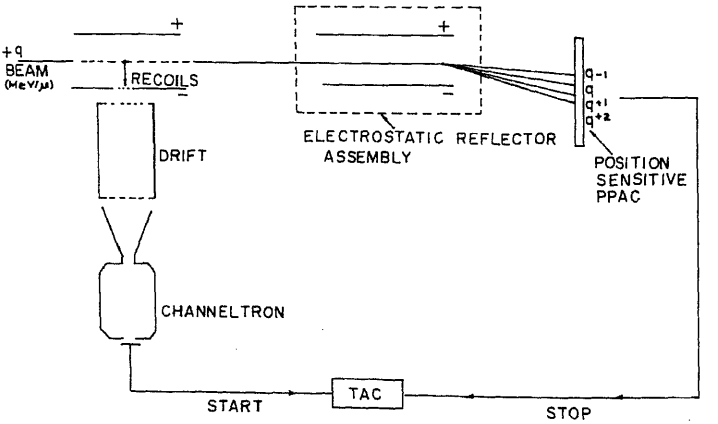


Figure 1. Schematic diagram of the set up for coincidences between recoils and projectile-ions.

h the neutral target atoms are separated using a TOF set up. The scattered projectiles ving in the forward direction are charge state separated by an electrostatic charge lyser and are finally detected by a one-dimensional position sensitive parallel plate lanche counter (PPAC). The fine control of the incident beam is facilitated by a four slit assembly. A brief description of the design and working principles of various nponents of the SCORPION set up is given below.

Four jaw slit assembly

e four jaw slit assembly connected to the exit port of the GPSC is an important part the SCORPION set up. It is required for controlling the incident number of rged particles because of the limited count rate handling capabilities of the PPAC 10^{4-5} particles/sec). It also helps in reducing the diameter of the incident beam ich is helpful in optimizing the time resolution of the recoil ion peaks as will be hlighted in § 2.2. The assembly is an octagonal aluminum block of thickness 100 mm, he centre of which a hole of 85 mm diameter is bored. The four orthogonal edges of s block are provided with a Wilson seal arrangement through which stainless steel S) shafts pass. The movement of the shafts is controlled with the help of micrometers. e rotational motion of the central shaft of the micrometer is converted into a linear tion of the SS shaft with the help of ball bearings. In order to lock even fine rotations the shaft during its linear motion, proper guides have been provided. Four SS strips, ted on the teflon pieces which are push fit at the ends of the shafts, form an aperture, size of which can be varied with the help of micrometers. In overlapping position, distance between any two strips is 2 mm. Each strip is connected to a hermetically led lemo connector for providing a current read out. The slit arrangement is lowed by a bellow (internal diameter (ID) $\sim$ 100 mm) and a cross (ID $\sim$ 100 mm). is combination houses a Faraday cup for the GPSC.

The exit port of the cross is connected to an SS scattering chamber through a gate ve. The chamber supported by a stand has an internal diameter of 300 mm and a height 150 mm. The TOF spectrometer is mounted in the chamber on the bottom port using an apter. The mounting arrangement has the facility of height adjustment of the ctrometer together with the possibility of its rotation and translation in a direction rpendicular to the beam direction. These features help in proper alignment of the ctrometer with respect to the incoming beam direction. One of the ports of the ttering chamber has a gas handling flange with appropriate mountings.

Time of flight spectrometer

e TOF set up separates the recoils of various charge states on the basis of their ne of flight over a certain path length. The set up consists of a pusher and an ractor plate (+ and −), a drift tube and a channeltron detector as shown in figure 1.

.1 *Time resolution of the TOF spectrometer*: The advantage of such a spectrometer is speed of collection of ions and a simple design, without any limitation on its shape and e. However, it is found to have a limited time resolution [11]. Ideally if all the recoil

ions formed in the interaction zone have zero initial velocity and move parallel to the extraction field, then the time of flight of all the ions having the same mass to charge ratio will be the same and in this case the resolution of the spectrometer will be limited by the resolution of the detector. In practice, the other limiting factors contributing to the time spread are:

2.2.1.1 *Initial space distribution of ions*: Due to a finite width of the beam and the finite spread of the gaseous target atoms, ions are born at different positions. This leads to a variation in flight times of ions having the same m/q ratio.

2.2.1.2 *Initial kinetic energy distribution of ions*: The time spread due to an initial kinetic energy distribution of the ions is due to the difference in flight times of two ions born at the same position but with the velocity vectors pointing in opposite directions.

2.2.2 *Estimation of time spreads*: If the ions are born with an initial energy U_0, then the total energy of the ions in charge state q impinging on the detector (see figure 1) is given by

$$U = U_0 + qsE_s + qdE_d + qxE_x \tag{2}$$

where E_s, E_d and E_x are the electric fields in the spaces s, d and x respectively. The total time of flight is given by the relation

$$T = T_s + T_d + T_D + T_x \tag{3}$$

where,

T_D = time of flight throught drift length D,

$$T_s = \frac{(2m)^{1/2}}{qE_s}[(U_0 + qsE_s)^{1/2} - U_0^{1/2}], \tag{4}$$

$$T_d = \frac{(2m)^{1/2}}{qE_d}[(U_0 + qsE_s + qdE_d)^{1/2} - (U_0 + qsE_s)^{1/2}], \tag{5}$$

$$T_D = \frac{(2m)^{1/2}D}{2(U_0 + qsE_s + qdE_d)^{1/2}}, \tag{6}$$

$$T_x = \frac{(2m)^{1/2}}{qE_x}[(U)^{1/2} - (U_0 + qsE_s + qdE_d)^{1/2}]. \tag{7}$$

As mentioned in § 2.2.1.1, the time resolution of the spectrometer is limited by the time spread due to initial space distribution of ions. To minimise this time spread, it is necessary that all the ions having the same m/q and born around the mean s_0 reach the detector at the same time. The condition is satisfied if, at the detector

$$\left(\frac{dT}{ds}\right)_{U_0 s_0} = 0. \tag{8}$$

Using condition (8) in (3) and assuming that the ions are born with zero initial velocity, we get

$$D = 2(sE_s + dE_d)^{3/2}\left[\left(\frac{1}{sE_s}\right)^{1/2}\left(\frac{1}{E_s}-\frac{1}{E_d}\right)+\left(\frac{1}{sE_s+dE_d}\right)^{1/2}\left(\frac{1}{E_d}-\frac{1}{E_x}\right)\right.$$
$$\left.+\left(\frac{1}{sE_s+dE_d+xE_x}\right)^{1/2}\left(\frac{1}{E_x}\right)\right], \tag{9}$$

where D gives the optimum drift length required for the best time resolution. To estimate the contribution from the time spread due to kinetic energy distribution (§ 2.2.1.2), consider two ions (ion 1 and ion 2) born at the same initial position s with the same initial speed but with their velocity vectors pointed towards + and − respectively. Ion 1 moves a certain length towards + before it is pulled back by the potential on −. Subsequently its motion is identical to that of ion 2 moving towards − from the beginning. Hence, the time difference between ion 1 and ion 2 reaching the detector is the 'turn around time' given by the relation,

$$T = \text{Turn around time}$$
$$= 2\,(\text{Deceleration time}) = 2\frac{(2mU_0)^{1/2}}{qE_s}. \tag{10}$$

2.2.3 *Mass resolution*: The maximum resolvable mass, M, [11] depends on the initial space and energy distribution functions. Consider the relation

$$\frac{1}{M} = \frac{1}{M_s} + \frac{1}{M_\theta}, \tag{11}$$

where M_s is the maximum resolvable mass for which time spread due to space distribution of ions ΔT_s equals the time spread between adjacent mass peaks and M_θ is the maximum resolvable mass for which the time spread due to kinetic energy distribution of ions ΔT_θ equals the time between adjacent mass peaks. M will be smaller than either M_s or M_θ. On the other hand, it will be at least as large as the value M obtained above by assuming the total spread to be the sum of the energy and space time spreads.

2.2.4 *The present TOF spectrometer*: The TOF spectrometer is a SS box of dimensions 100 mm × 60 mm × 100 mm. The spectrometer, besides housing the TOF set up, has provisions for

(1) adjustable collimators for beam entrance and exit made from tantalum.
(2) arrangement for gas input through a hypodermic needle (diameter = 1 mm) with a provision for adjusting its height with respect to the incoming beam. The needle is connected to the gas handling flange through a nalgene tube of diameter 10 mm. The gas flow in the interaction region is controlled with the help of fine needle valves.

+ and − are copper plates of dimensions 20 mm by 20 mm and thickness 4 mm. The plates are separated from each other by 20 mm (= $2s$) with the help of teflon spacers. The extractor plate has a 5 mm circular hole in the centre for allowing the recoils to pass through after extraction from the interaction region. To avoid field penetration through the hole and through the ends of the drift tube, a grid of parallel wires (25 micron diameter wires spaced 1 mm apart) is epoxied using a conducting glue. The transmission

efficiency of the grid is about 90%. The 30 mm ($= D$) long drift tube is separated from the extractor plate by 3 mm ($= d$). Following the drift tube is the channeltron detector spaced 5 mm ($= x$) from the end of the former.

2.3 *Parallel plate electrostatic charge deflector*

An ion of mass m, having an energy T and a charge state q suffers a deflection X when it passes through a pair of plates across which an electric field of E (kV/cm) exists. The deflection X is given by the expression

$$X = \frac{qeE}{4T}[L(L+2d)], \tag{12}$$

where L is the length of the deflector plate and d is the separation between the end of the deflector plate and the detector for the ions.

Using the above principle, a parallel plate electrostatic charge analyser set up was fabricated and mounted at the exit port of the scattering chamber. The assembly consists of two parallel aluminum plates, each 400 mm long and 60 mm wide. The plates are housed in a 150 mm O.D. and 600 mm long beam pipe. High voltage is applied to the plates through the specially designed feedthru's which are detachable. The feedthru is insulated from the beam pipe using teflon cylinders which pass through a Wilson seal arrangement provided in the housing for the plates. The length of the trapped air column between the feedthru and the teflon cylinder is reduced using a single O-ring which is pressed onto the plate by the cylinder. Such an arrangement reduces the chances of breakdown because of poor vacuum due to trapped gases. The place of application of high voltage is covered by a teflon cap. The gap between the plates can be varied with the help of nylon spacers. In the present arrangement, the plates are separated by 25 mm with the top plate being 5 mm above the beam height. The parallel plate deflector is followed by a bellow, a long tube, a TEE, a gatevalve and a PPAC. The total flight path of the ions after exiting from the charge analyser to the PPAC is 1600 mm, which is sufficient to provide a good separation between the charge states of the scattered projectile. The bellow helps in bending the entire assembly in such a way that for a given electric field between the plates, the incident charge state hits the centre of the detector with the charge states due to electron capture and loss on either side. The expected deflection of a 60 MeV Si^{4+} ion for a field of 1.2 kV/mm between the plates, as found out using (12) is 29 mm and the separation between neighbouring charge states is 7 mm.

The SCORPION incorporates pumping stations which are fitted with turbomolecular pumps to provide a oil-free vacuum.

2.4 *One dimensional position sensitive PPAC*

A one dimensional position sensitive PPAC [12] detector, mounted at the end of the SCORPION, is housed in a SS flange of diameter 100 mm. The window foil of the detector is a 1.6 μm aluminized mylar foil. The anode is a similar foil glued to 1.6 mm thick PCB. The cathode is a printed circuit board having aluminized parallel strips. The width of each strip is 1.4 mm and the spacing between the strips is 0.6 mm. The cathode is made position sensitive using the resistive chain method.

The value of the resistance between any two strips is 300 Ω. The spacing between two electrodes is 5 mm. The anode is kept at a high voltage whereas the two ends of the cathode are grounded through 100 MΩ resistances. The active area of the PPAC is 30 mm × 70 mm.

3. Commissioning tests of SCORPION

The entire assembly was pumped to a vacuum of 4×10^{-6} mbar. The GPSC, pumped by a 2000 l/sec diffstak pump showed a vacuum of 1×10^{-6} mbar. For initial tests, the PPAC was replaced by a glass viewport with quartz pieces stuck along the vertical central axis. With the slits fully open, a beam of 60 MeV Si^{4+} ions was allowed to fall onto alumina pieces stuck to the sides of the entrance collimator of the TOF spectrometer. The beam was focussed and steered so as to pass through the tantalum collimator. With a suitable electric field applied to the deflector plates, the bent beam was viewed on the quartz pieces of the glass ports. With changes in the voltages applied to the plates, the movement of the beam on the quartz pieces was observed. The tests ensured correct alignment of all the components of the SCORPION set up. The slits were closed using the micrometers and the decrease in the intensity of the spot was observed as expected.

3.1 *Coincidence measurements between recoils and scattered projectiles*

For the coincidence measurements, the view port at the end of the assembly was removed and replaced by the PPAC. Initially, the gate valve isolating the PPAC from the rest of SCORPION line was closed and the slits were fully opened. A pulsed, 60 MeV Si^{4+} beam from the accelerator, with a repetition rate of 1 μs, was properly focussed and passed through the spectrometer so as to crossfire a neutral beam of argon atoms. The flow of the gas was controlled with the help of the needle valves. The recoils in various charge state were extracted from the interaction region and made to drift towards the channeltron detector by using optimized voltages on the different components of the set up, which are listed in table 1. The signals from the channeltron were fed into a fast, current sensitive preamplifier. The TOF spectrum of the recoils was obtained by giving the 'start' signal to the time-to-amplitude converter (TAC) from the channeltron and a 'stop' signal from the rate divided RF. Single collision conditions

Table 1. Optimized values of voltages and distances in the time of flight set up.

Distance of the interaction region from − (s)	10 mm
Distance of the drift tube from − (d)	3 mm
Length of the drift tube (D)	30 mm
Distance of CEM from the drift tube (x)	5 mm
Voltage on + plate	+1 kV
Voltage on − plate	−1 kV
Voltage on the drift tube	−1.3 kV
Voltage on the channeltron	−2.2 kV

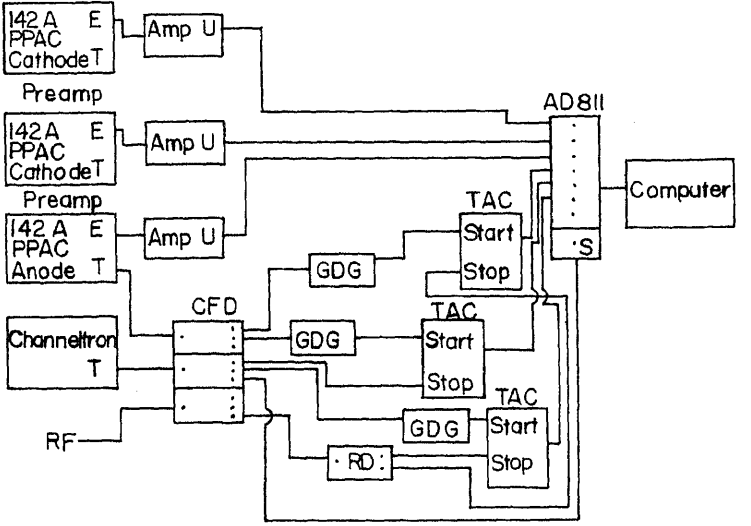


Figure 2. Schematic diagram of the electronics used to process the signals for the SCORPION experiment.

were ascertained by studying the variation in the yield of Ar^+ recoil ions normalized to the number of ions detected by the PPAC as a function of the gas pressure. In the absence of a capacitance manometer, the pressures were measured at the gas entrance with the help of a Pirani gauge and just below the TOF spectrometer with the help of a Penning gauge. After ensuring that single collision conditions prevailed in the interaction region, fine tuning of the beam was done to get an optimized argon recoil spectrum. The observed count rate of the channeltron was about 3000 counts per second. Following this the slits on the four jaw assembly were closed in a controlled fashion with the TOF spectrum being monitored after each change. Maintaining a consistent recoil-ion spectrum, the count rate in the channeltron was reduced from 3000 counts per second to 10 counts per second. Subsequently, voltages of ± 10 kV were applied to the deflector plates and the gate valve isolating the PPAC from the rest of assembly was opened. Slits were fine tuned so that the PPAC counted at a rate of 8 kHz to 10 kHz.

The electronic block diagram for processing the signals from the channeltron, the PPAC anode and the two ends of the PPAC cathode is shown in figure 2. The processed signals were fed to an analog to digital converter (ADC). A data acquisition program NSCSORT, is used to acquire the data in list mode. The strobe to the ADC was obtained from the channeltron. A typical TOF spectrum of the recoils produced in 60 MeV Si^{4+}–Ar collisions recorded from the channeltron–PPAC TAC is shown in figure 3 without any gating from the charge selected projectiles. The position spectrum of the charge separated projectiles shown in figure 4 for the 60 MeV Si^{4+}–Ar case was obtained by dividing the energy signal from one end of the cathode with the sum of the signals from both the ends. In addition to the incident beam charge state, projectiles which have lost up to 4 electrons are also seen. With the change in voltage of ± 2 kV on the deflector plates, the change expected in the position of the projectile peaks is ~ 4 mm. A

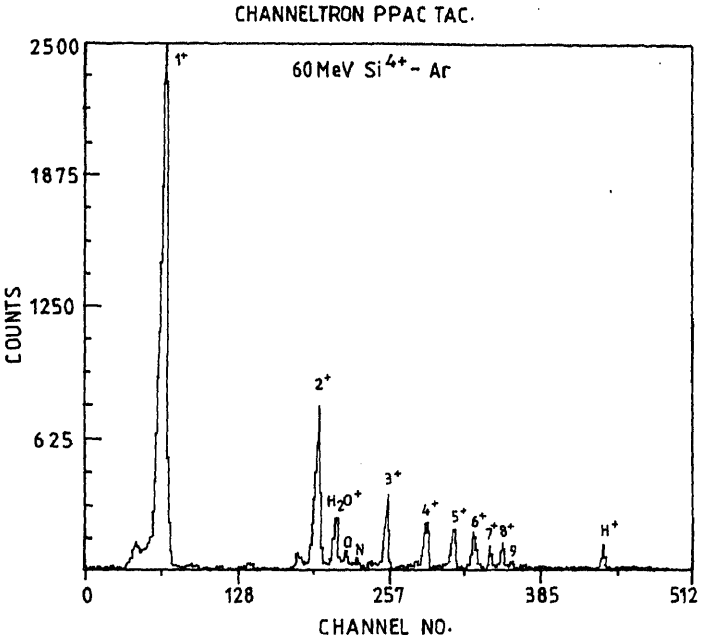


Figure 3. Time-of-flight spectrum of recoils obtained from the channeltron-PPAC-TAC for 60 MeV Si^{4+} ions interacting with argon gas.

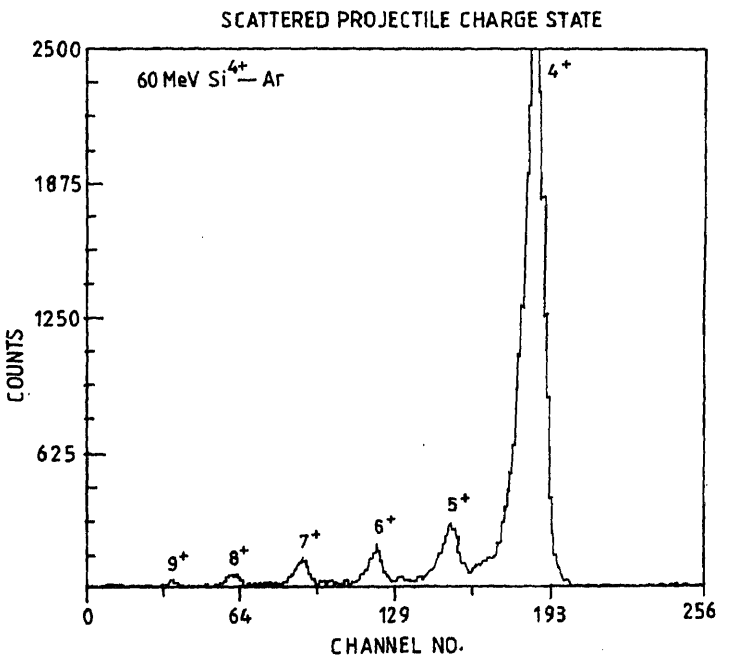


Figure 4. Position spectrum of the scattered projectiles produced in 60 MeV Si^{4+}–Ar collisions.

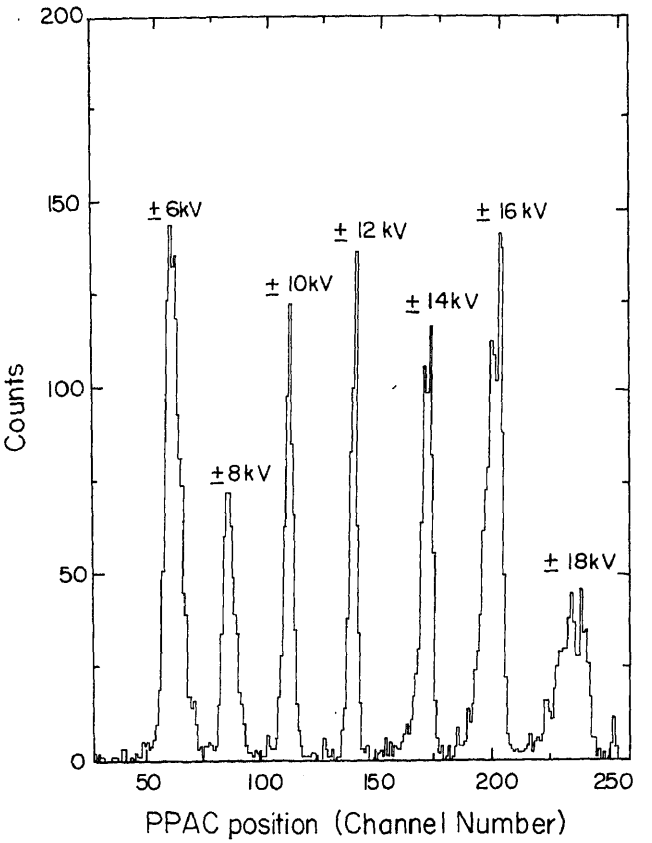


Figure 5. Position spectrum of the 60 MeV Si^{4+} charge state of the beam for the increase in the deflector plate voltage in steps of ± 2 kV.

typical position spectrum of the 60 MeV Si^{4+} charge state of the beam, for the variation in the voltages on the deflector plates in steps of ± 2 kV, is shown in figure 5. It may be noted that the time of aquisition is not the same for all the peaks. The channel numbers along the X axis correspond to the position of the PPAC in the vertical direction along which the deflection is observed.

4. Recoil ion distributions for various ionization channels

The recorded data were processed using the NSCSORT program in order to extract recoil ion spectra corresponding to particular charge states of the projectiles. With conditions applied to the total recoil-ion spectrum, the recoil ion distributions for direct ionization and for electron loss up to 4 electrons were obtained as shown in figures 6(a–e). For the direct ionization process (figure 6a), the recoil ion yield drops with the increase in the recoil charge state, indicating collisions with large impact parameters. The recoil ion distributions become bell shaped for the case of

electron loss (figure 6b–e), with the peak of the distribution shifting to higher recoil charge states. This bell shaped distributions point to collisions of smaller impact parameters as compared to the direct ionization. The shifting of the peak to the higher recoil charge state with the increasing number of electrons lost by the incident ion indicates a deeper penetration by the projectile [3]. Note that for Si^{4+} incident beam, electron capture processes are not observed. On increasing the charge state of the incoming beam, the position spectrum of the scattered projectiles shows peaks which correspond to electron capture processes in addition to direct ionization and electron loss processes. Detailed results on the studies of direct ionization, electron capture and electron loss processes and their effect on the recoil ion distributions for 60–120 MeV Si^{q+}–Ar collision system will be reported elsewhere in a separate publication.

5. Conclusion

A facility to perform coincidences between recoil ions produced in fast, heavy, ion-atom collisions with the charge separated projectiles, has been developed and installed at NSC, New Delhi. The design features of the various components have been described in detail. This facility has been used to study the direct ionization, electron capture and electron loss channels in collisions of multiply charged silicon ions with neutral argon atoms. Some typical spectra of the recoil ions for various channels have been presented and briefly discussed.

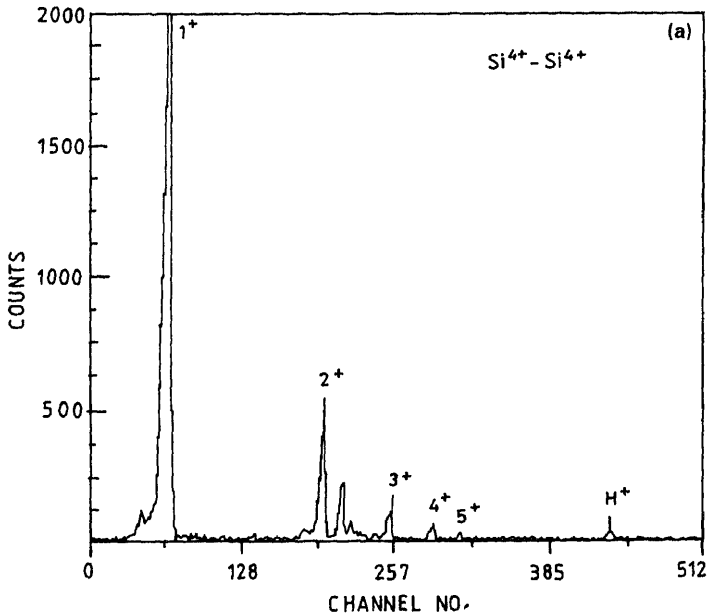


Figure 6(a).

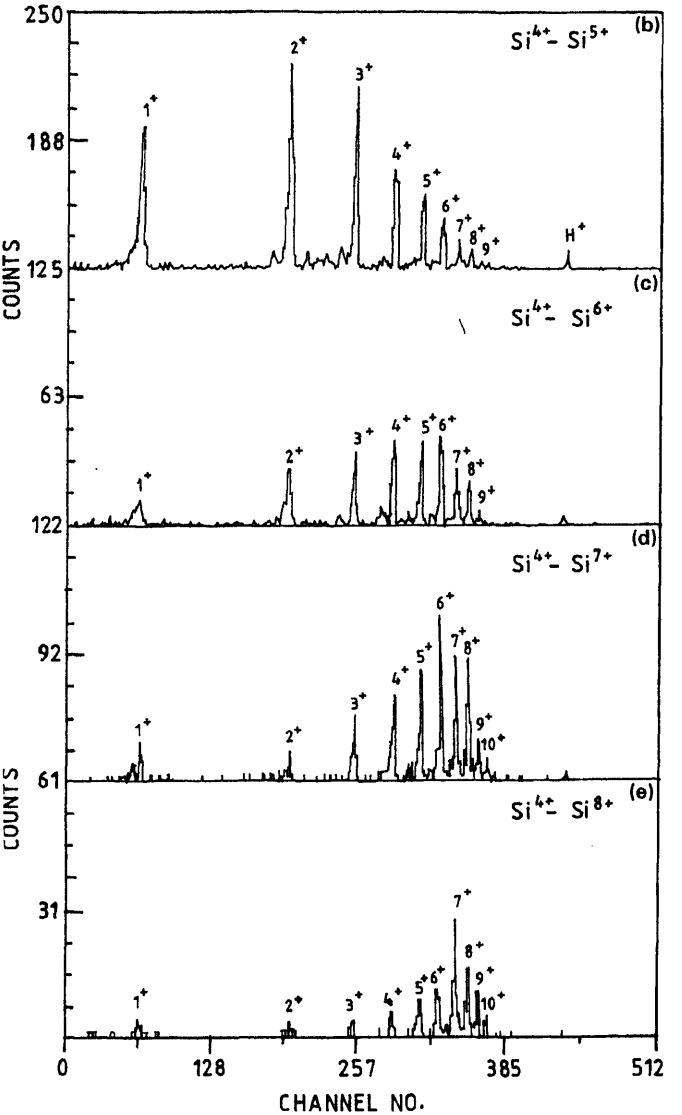


Figure 6. Recoil distributions corresponding to (**a**) direct ionization, (**b**) 1-e^- loss, (**c**) 2-e^- loss, (**d**) 3-e^- loss and (**e**) 4-e^- loss events in the case of 60 MeV Si^{4+}–Ar collisions.

Acknowledgements

The authors acknowledge the constant encouragement, support and guidance of Prof. G K Mehta. Useful suggestions from E Krishnakumar, R Dörner and T Tonuma at the design stages of this work are acknowledged. The authors are grateful to J Chacko for precise and excellent machining of the various components of the set up and to Sunder Rao for excellent welding jobs. The concerted effort put in by the pelletron personnel and the

members of the pulsing group for delivering high quality beams during the experiment is acknowledged.

References

[1] T J Gray, C L Cocke and E Justiniano, *Phys. Rev.* **A22**, 849 (1980)
[2] A Muller, B Schuch, W Groh, E Salzborn, H F Beyer, P H Mokler and R E Olson, *Phys. Rev.* **A33**, 3010 (1986)
[3] T Tonuma, H Kumagai, T Matsuo and H Tawara, *Phys. Rev.* **A40**, 6238 (1989)
[4] R Parameswaran, R L Watson, V Horvat, G Sampoll and D A Church, *Z. Phys.* **D32**, 61 (1994)
[5] C L Cocke, *Phys. Rev.* **A20**, 749 (1979)
[6] G K Mehta and A P Patro, *Nucl. Instrum. Methods* **A268**, 334 (1988)
[7] M J Singh, S K Goel, R Shanker, D O Kataria, N Madhavan, J J Das, A Tripathi, P Sugathan, D K Avasthi and A K Sinha, *Indian J. Phys.* **B69**, 181 (1995)
[8] J Ullrich, K Bethge, S Kelbch, W Schadt, H Schmidt Bocking and K E Steibing, *J. Phys.* **B19**, 437 (1986)
[9] T Tonuma, T Matsuo, H Kumagai, S H Be and H Tawara, *RIKEN Accel. Prog. Rep.* **18**, 197 (1984)
[10] D K Avasthi, Ambuj Tripathi, D Kabiraj, S Venketaramanan and S K Datta, *NSC Annual Report* **30**, 42 (1991)
[11] W C Wiley and I H Mclaren, *Rev. Sci. Instrum.* **26**, 1150 (1955)
[12] J J Das, D O Kataria, Jaipal, N Madhavan and A K Sinha, *DAE Symp. Nucl. Phys.* **B36**, 412 (1993)

—journal of
physics

November 1997
pp. 535–546

TDPAC measurements in textured hcp samples

R GOVINDARAJ and K P GOPINATHAN*
Materials Science Division, Indira Gandhi Centre for Atomic Research, Kalpakkam 603 102, India
* Department of Physics, Cochin University of Science and Technology, Cochin 682 022, India

MS received 28 July 1997; revised 7 October 1997

Abstract. Time differential perturbed angular correlation measurements done on the 482 keV level of ^{181}Ta probe nuclei in well annealed reference foil samples of α-ZrHf, α-TiHf and α-Hf indicate that all probe nuclei are defect free, substitutional and experience quadrupole frequencies characteristic of these matrices. But mostly the ratio of the intensities of the harmonic components of the quadrupole frequency is never in 3 : 2 : 1, which is understood to be due to the selective orientation of microcrystallites brought out by the pre thermo-mechanical treatments on the sample. This paper illustrates as to how texturing in general could cause ambiguities in the measurements of the hyperfine interaction parameters of probe nuclei particularly in hcp foil samples with or without the association of any defects and the ways to delineate the effect of texturing.

Keywords. TDPAC; electric field gradient; texturing.

PACS Nos 81.40; 82.20

1. Introduction

The hcp samples like α-Hf, α-Ti and α-Zr are non-magnetic. The quadrupole interaction (i.e., the interaction between the quadrupole moment of probe nuclei and electric field gradient at their sites) at the isomeric state separating 133 and 482 keV γ–γ cascade of ^{181}Ta probe nuclei in these samples leads to the partial removal of degeneracy of the isomeric state into doubly degenerate states with m-values $\pm 5/2$, $\pm 3/2$ and $\pm 1/2$. Choosing the direction of the first γ ray as the quantization axis leads to the selective population of the above magnetic substates. This results in the emission of the second γ ray of the above cascade of the probe nucleus with an altered population distribution and this change is responsible for the attenuation of the angular correlation [1, 2]. The Fourier transform of the anisotropy spectra in the case of the probe nuclei occupying defect free axially symmetric charge environment in a polycrystalline hcp sample having microcrystallites oriented randomly, should have peaks at ω_1, ω_2 and ω_3 having the ratio of 1 : 2 : 3 and the corresponding intensity ratio $\approx 3 : 2 : 1$. The symmetry of the charges that are distributed at the nearest neighbouring environment of probe nuclei is characterized by an asymmetry parameter which is roughly given as $(2 - \omega_2/\omega_1)$.

The probe nuclei occupying the defect free environment of the above matrix in a single crystal of a non-cubic and axially symmetric sample will experience the same quadrupole frequency having the similar ratio between its harmonic components as that of the polycrystalline sample. However the ratio of the intensities of the components of the

quadrupole frequency will not be 3 : 2 : 1 and is sensitively dependent upon the orientation of the crystal with respect to detector geometry. Measurements in a single crystal indicate a drastic difference in intensities of the components of the quadrupole frequencies for different orientations of the sample with respect to detector geometry. By determining the orientation of the respective principle axes of the EFG's we can find out the geometry of the defects as trapped by probe nuclei in the measurement [3]. A textured sample is essentially composed of preferentially oriented and random microcrystallites with the former out numbering the latter. Therefore in the case of the textured sample for some fraction of the probe nuclei the principal axis of the EFG tensor will be oriented selectively and for the remaining fraction the latter will be in random. This carries the information about the orientation of microcrystallites in a sample. The formulation of angular correlation and distribution experiments in the presence of a texturized hyperfine interaction has been given by Rots *et al* [4].

The texturing effects if present *predominantly* in a hcp sample, could cause ambiguities regarding the hyperfine interaction parameters as experienced by probe nuclei viz., quadrupole frequency and asymmetry parameter. The former could be the case if the intensity of the fundamental component of the quadrupole frequency is almost zero. The latter is possible if the intensities of the second and third harmonic components of the quadrupole frequency are almost zero which in general would give an impression that the probe nuclei are occupying an axially asymmetric charge environment. These effects have to be seriously considered in textured samples in which defects are introduced by any conventional method such as light or heavy ion implantation, quenching or due to the introduction of extrinsic defects such as impurities. All TDPAC measurements in such a sample following each step of any heat treatment, in general has to be necessarily carried out in an unique position of the sample with respect to detector geometry. This is required to be done since the anisotropy spectra and hence its Fourier transform would be varying in terms of the intensities of the frequency components corresponding to probe nuclei occupying oriented microcrystallites.

Considering the case of a hcp sample, if it is well homogenized and in the form of a spherical lump it would most likely be perfectly polycrystalline with a random orientation of microcrystallites. Therefore the probe nuclei occupying the substitutional and defect free environment in it will have the ideal intensity ratio of the frequency component corresponding to the concerned matrix. Even such samples under certain conditions seem to have texturing behaviour. For example the preferred orientation of microcrystallites has been reported in In metal just below the melting point by ^{111}Cd TDPAC measurements [5]. Rasera *et al* [6] have proposed a method of rotation of the sample around an axis perpendicular to the plane of the detectors to get a polycrystalline analogue of the texturized samples.

In this paper we are aimed at the following: (1) the illustration of *texturing* effects in hcp foil samples and the possible ambiguities that arise in resolving and/or interpreting the intrinsic or extrinsic defects in such samples, (2) to experiment the ways to get the quadrupole interaction parameters analogue of the polycrystalline sample with random microcrystallites and hence resolve the hyperfine interaction parameters of probe nuclei associated with defects if any occur in the sample. Herewith we have proposed a method to obtain the polycrystalline analogue of the textured sample.

The TDPAC of the 133–482 keV γ–γ cascade of ^{181}Ta was measured in the samples of Hf, ZrHf and TiHf samples respectively by a three detector twin fast-slow coincidence setup [7] having NaI(Tl) detectors. One of the detectors was gated for the START (133 keV) γ-ray while the other two detectors were located at 90° and 180° respectively with respect to the START detector to detect the STOP (482 keV) γ-ray. The delayed time resolution spectra were obtained in the form of the count rate as a function of the time elapsed after the emission of the first γ-ray. The two time spectra $W(90°, t)$ and $W(180°, t)$ were recorded simultaneously. The prompt time resolution of the setup measured with a ^{60}Co source was 2.2 ns FWHM when gated for the above cascade of ^{181}Ta [7].

From the delayed time resolution spectra $W(90°, t)$ and $W(180°, t)$, the normalized anisotropy function $R(t)$ was calculated as

$$R(t) = A_2 G_2(t) = 2 \times \frac{[W(180°, t) - W(90°, t)]}{[W(180°, t) + 2W(90°, t)]}, \tag{1}$$

where A_2 is the effective anisotropy coefficient and $G_2(t)$ is the perturbation factor. $W(90°, t)$ and $W(180°, t)$ are the coincidence count rates at 90° and 180° respectively between the START and STOP detectors. In general during the data acquisition using a three detector PAC system the START detector is moved at periodic intervals between 90° and 180° positions with respect to either of the STOP detector and the delayed time resolution spectra are obtained in the form of the count rate as a function of the time elapsed after the emission of the first γ-ray [8]. Hence for each position of the START detector two time spectra $W(90°, t)$ and $W(180°, t)$ are recorded simultaneously. This is done essentially to cancel out the term representing the efficiency of the detectors that appears in eq. (1), so that one gets the normalized anisotropy coefficient independent of any detector parameter. In this work some experiments are done as prescribed above and the remaining with a single geometry of the detecting system i.e. by keeping the START detector fixed. The latter method is justified because of the reasoning that two STOP detectors are identical and their detection efficiencies remain almost the same.

The $R(t)$ spectra were least squares fitted to the function [2]

$$R(t) = A_2 \sum_{i=1}^{r} G_2^i(t), \tag{2}$$

$$A_2 G_2^i(t) = A_2 \sum_{m=0}^{3} a_m^i \exp[-\delta_i k_m^i(\eta_i)\omega_{Q_i} t] \cos[k_m^i(\eta_i)\omega_{Q_i} t], \tag{3}$$

where $k_0^i = 0$, $k_1^i(\eta_i) + k_2^i(\eta_i) = k_3^i(\eta_i)$ and $\sum_{m=0}^{3} a_m^i(\eta_i) = 1$.

If the probe nuclei occupy axially symmetric aand non-cubic sites in a sample in which microcrystallites are randomly oriented the values of the amplitudes a_m^i are given as [2]

$$a_m = \sum_{m,m_1}^{I} \begin{pmatrix} I & I & 2 \\ m_1 & -m & m - m_1 \end{pmatrix}, \tag{4}$$

where $2|m^2 - m_1^2| = n$ for half-integer I. The values of a_m are tabulated for the isomeric state of different spin as given in [2]. In the case of a textured sample these a_m^i have to be taken as free parameters while doing the data analysis.

The spin value of the isomeric state of ^{181}Ta being 5/2, we have

$$\langle \nu_Q \rangle = eQV_{zz}/h = 10\omega_Q/3\pi, \tag{5}$$

where V_{zz} is the principal component of the electric field gradient (EFG) tensor. When the EFG is not axially symmetric, the asymmetry parameter $\eta = (V_{xx} - V_{yy})/V_{zz}$ where $|V_{zz}| \geq |V_{yy}| \geq |V_{xx}|$ is extracted from the fit of $R(t)$ data to equation (2). The value of r which occurs in (2) is dependent upon the number of frequency components that appear in the Fourier transform of the anisotropy spectra.

The EFG tensor is completely determined by the frequency ν_Q and the asymmetry parameter η. In the case of probe nuclei being present in a non-cubic and axially charge symmetric environment in a polycrystalline sample the value of the asymmetry parameter will be zero which will result in $k_m(\eta = 0) = m$. The probe nuclei present in an asymmetric charge environment are characterized by an asymmetry parameter which is given as $\eta \approx 2 - (k_2/k_1)$. The exponential factor in the sum accounts for the possible existence of a quadrupole frequency distribution which is assumed to be correctly described by a Lorentzian shape with a relative width δ. A non-vanishing value of δ implies either a significant concentration of defects and/or impurities in the material under study, or with a disordered arrangement of atoms in the probe surroundings [8, 9].

The ratio of $a_1 : a_2 : a_3$ will be approximately 3 : 2 : 1 in the case of the probe nuclei present in an axially charge symmetric non-cubic sample in which the microcrystallites are randomly oriented. The probe nuclei present in a single crystal of the above sample will experience the same quadrupole frequency as that of the polycrystalline sample. But the difference between these two cases is shown up in the values of $a_m(\eta)$ which are dependent upon the orientation of the crystal and hence the direction of the principal axis of the EFG tensor with respect to detector geometry [10]. The theoretical expression for the time dependent anisotropy spectra $R(t)$ in a textured sample can in principle be written as a sum of fractions of probe nuclei experiencing an EFG whose principal component is oriented preferentially along some direction and the remaining fraction with the latter oriented randomly [11]. A detailed XRD measurement has to be done to get any quantitative information on texturing of the sample. But selective orientation of impurity precipitates if any in a sample could be found out by TDPAC, where XRD cannot be used if the concentration of impurity is less than the detectable limit.

3. Results and discussion

We shall first consider the results of the TDPAC measurements carried out on helium implanated α-Zr sample [12]. The homogeneous helium implanation of the sample of dimension 1 cm×0.1 cm×380 μm has been carried out with a 40 MeV alpha beam to a dose of 100 appm at VEC, Calcutta. Measurement on α-ZrHf reference sample has shown that all probe nuclei experience a Lorentzian distribution of quadrupole frequencies having a value of $\langle \nu_{Q_0} \rangle = 310 \pm 4$ MHz and an asymmetry parameter $\eta_0 = 0.08$. This is

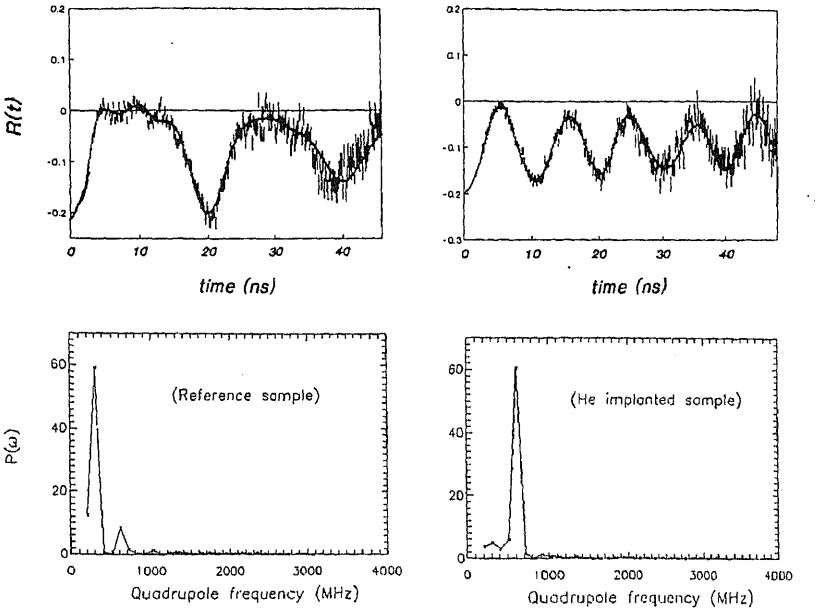


Figure 1. The experimental $R(t)$ vs t spectra and its Fourier transformed spectra at room temperature in the reference and He implanted α-ZrHf sample. The continuous curve in $R(t)$ is the one calculated using the fitted values of the parameters.

in agreement with the results reported in the literature [13]. The $R(t)$ spectra and its Fourier transform of the helium implanted sample in principle could lead to a misleading inference that a large fraction (≈ 0.89) of probe nuclei are associated with a unique helium implantation induced defect experiencing a quadrupole frequency of 620 MHz and the remaining are defect free and substitutional. As this frequency coincides with the second harmonic component of the quadrupole frequency corresponding to the substitutional fraction of probe nuclei we have to find out whether the observed effect is due to texturing or due to any defect trapping by a major fraction of probe nuclei. This can be found out by carrying out the measurements in different orientations of the sample with respect to detector geometry. The variation of the intensities of the quadrupole frequency components with different orientations of the sample confirms that the result is due to the texturing effects of the sample and not due to probe nuclei trapping any defects.

In the present case the selective orientation of the microcrystallites has resulted in a larger intensity of the second harmonic component of the quadrupole frequency as observed by the probe nuclei shown in figure 1. Data analysis of the $R(t)$ spectra on the helium implanted sample taking into consideration the texturing effects of the sample has shown that the probe nuclei experience the following hyperfine interaction parameters viz., $\langle \nu_{Q_1} \rangle = 310 \pm 4$ and $\eta_1 = 0.10$. Above results indicate that all probe nuclei occupy substitutional sites in hcp Zr matrix as reported in the literature [13]. A low value of the asymmetry parameter η in the above cases implies an almost axially charge symmetric environment of probe nuclei.

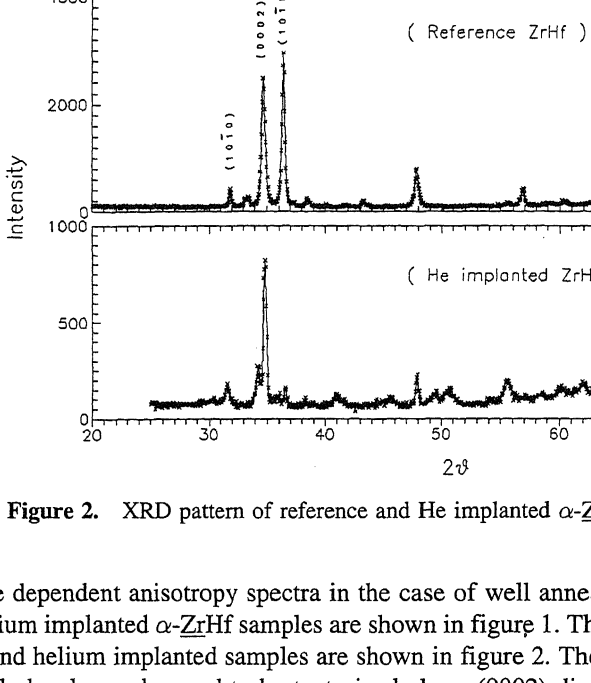


Figure 2. XRD pattern of reference and He implanted α-Z

The time dependent anisotropy spectra in the case of well anne and the helium implanted α-ZrHf samples are shown in figure 1. Th reference and helium implanted samples are shown in figure 2. Th ZrHf sample has been observed to be texturized along (0002) dire texturing effect as has been observed in the helium implanted samp due to the pre-thermo mechanical treatments of the sample. This m implanted α-Zr-Hf sample illustrates as to how texturing effects o defects (intrinsic or extrinsic) are likely to be occuring, coul misinterpretation about determining the hyperfine parameters of de associated fraction(s) of probe nuclei.

The effect of texturing is further illustrated by the results o measurements on Hf foil. The anisotropy spectra $R(t)$ correspondin in the Hf foil sample of area 1 cm$\times$0.1 cm and 25 μm thickne positions of the sample viz., the flat surface of the sample lie

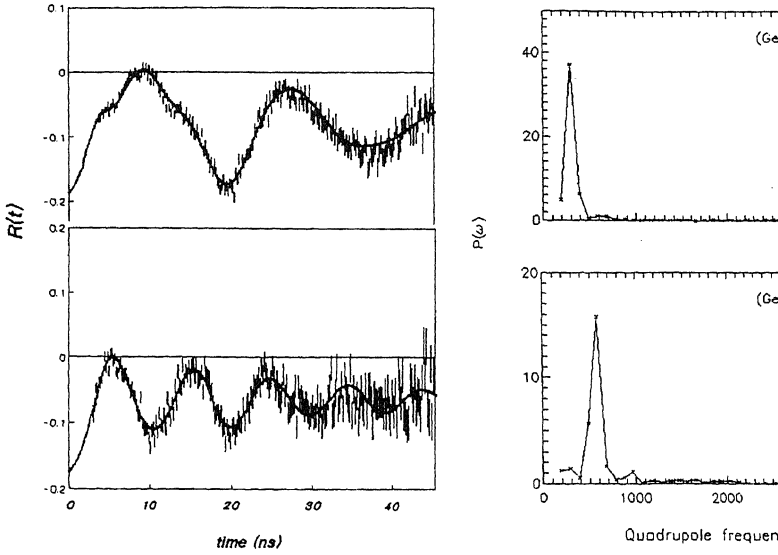
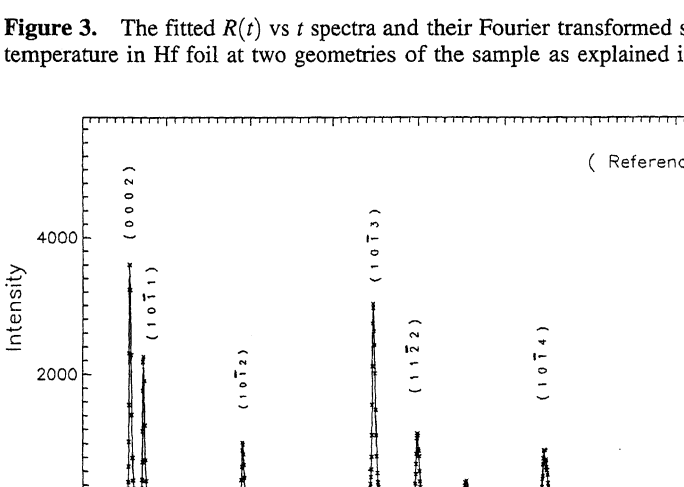


Figure 3. The fitted $R(t)$ vs t spectra and their Fourier transformed s
temperature in Hf foil at two geometries of the sample as explained i



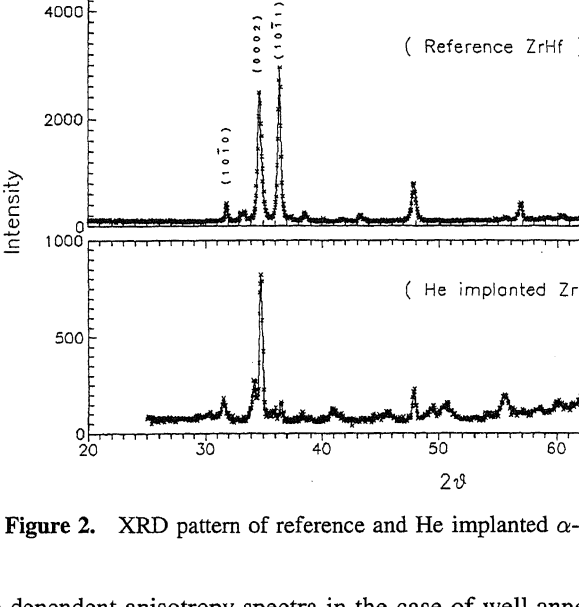


Figure 2. XRD pattern of reference and He implanted α-

The time dependent anisotropy spectra in the case of well anne and the helium implanted α-ZrHf samples are shown in figure 1. T reference and helium implanted samples are shown in figure 2. Th ZrHf sample has been observed to be texturized along (0002) dir texturing effect as has been observed in the helium implanted sam due to the pre-thermo mechanical treatments of the sample. This i implanted α-Zr-Hf sample illustrates as to how texturing effects c defects (intrinsic or extrinsic) are likely to be occuring, cou misinterpretation about determining the hyperfine parameters of d associated fraction(s) of probe nuclei.

The effect of texturing is further illustrated by the results c measurements on Hf foil. The anisotropy spectra $R(t)$ correspondi in the Hf foil sample of area 1 cm$\times$0.1 cm and 25 μm thickn positions of the sample viz., the flat surface of the sample li

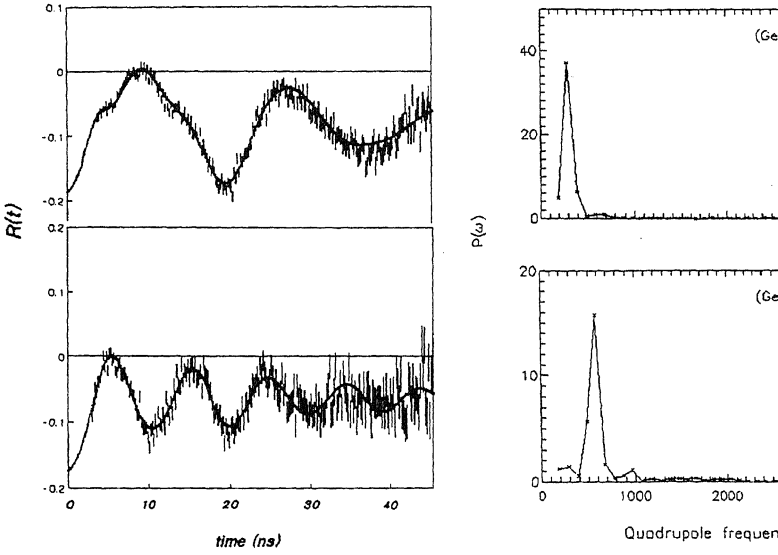
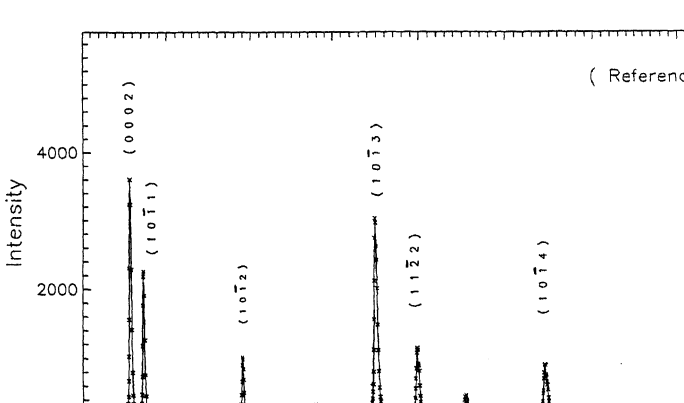


Figure 3. The fitted $R(t)$ vs t spectra and their Fourier transformed s
temperature in Hf foil at two geometries of the sample as explained i



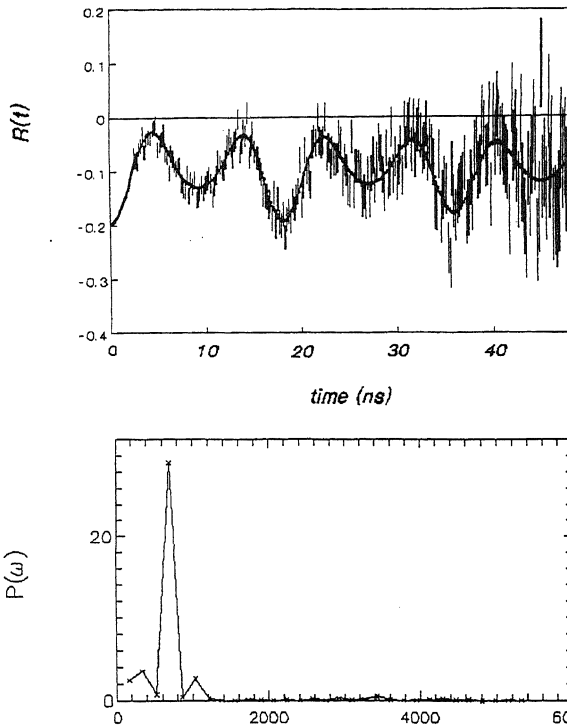


Figure 5. The fitted $R(t)$ vs t spectra and their Fourier tra
temperature in TiHf cold worked sample following the ann

frequencies with a relative width $\delta = 0.042$ having a mean valu
and an asymmetry parameter $\eta = 0.04$. This agrees with the repor
nuclei occupying defect free substitutional sites in hcp Ti matrix.
its Fourier transform following isochronal annealing treatment
which is nearly the re-crystallization temperature of Ti, implies th

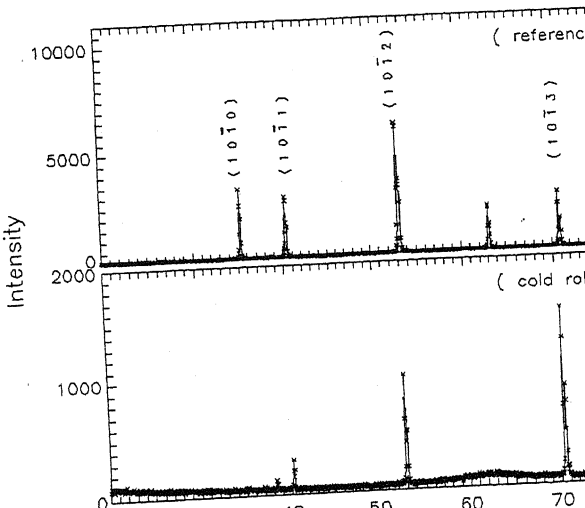


Figure 6. XRD pattern for reference, cold worked and is
sample.

to be taken to keep the sample at an unique position with resp
during the course of the measurements. Precipitation of probe i
observed by TDPAC measurements in any sample. In some case
get oriented preferentially under certain thermal treatments. We
orientation of Hf precipitates in Cu, Al matrix by ^{181}Ta TDPA
All these measurements in α-Hf, α-ZrHf and α-TiHf illustra
these samples.

We have proceeded in the following way the measurements
the effects of texturing and hence get the analogue of the perfec
and also the defect(s) related hyperfine parameter(s) of probe

(1) In this method, for getting the polycrystalline analogue
have proceeded as follows. We have averaged the experi

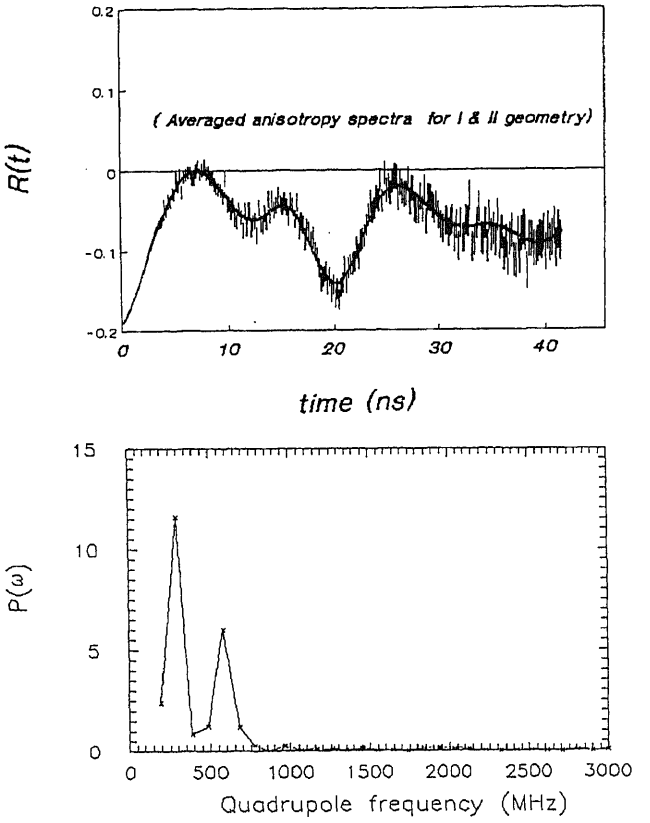


Figure 7. The fitted $R(t)$ vs t spectra and their Fourier transformed spectra at room temperature in Hf foil. The $R(t)$ is the average of the experimental spectra obtained at two geometries of the Hf sample as explained in the text.

of the quadrupole frequency components is seen with increasing speed of rotation of the sample. This has been illustrated in figure 8. The probe nuclei are found to experience a quadrupole frequency $\langle \nu_{Q_0} \rangle = 290 \pm 4$ MHz and an asymmetry parameter $\eta_0 = 0.08$. By rotating the sample about an axis perpendicular to the plane of the detector, we could average the selective orientations of microcrystallites of the sample. Therefore we may infer that the fraction of probe nuclei occupying the oriented microcrystallites may not contribute for the total anisotropy observed in the experiment. Hence this leads to a slight loss of anisotropy depending upon the speed of rotation. This can be seen in figure 8.

The measurements conducted either by averaging the $R(t)$ obtained at different geometries of the sample or by rotating it at a constant speed of rotation during the measurements would result in the polycrystalline analogue of the textured samples. These measurements would help us to find out the hyperfine interaction parameters of the defect(s) related fraction(s) occurring in such hcp samples without any ambiguity as we are able to delineate the fraction of probe nuclei that are defect free in the matrix.

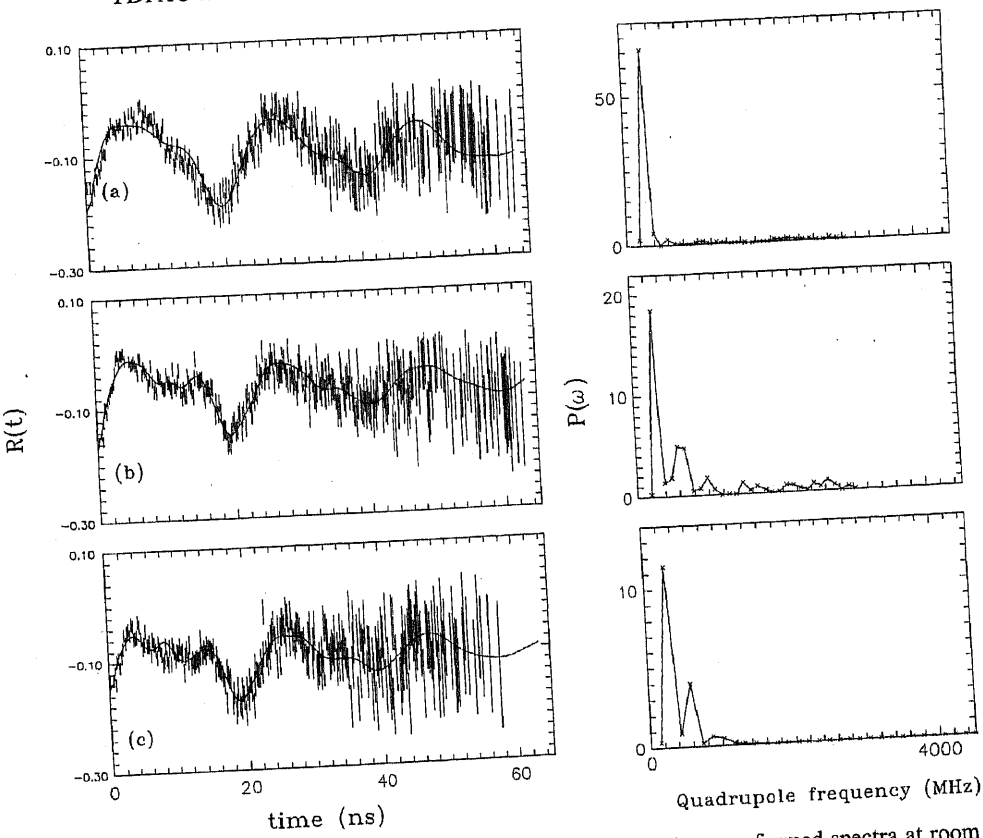


Figure 8. The fitted $R(t)$ vs t spectra and their Fourier transformed spectra at room temperature in Hf foil under the following conditions: **(a)** without rotation, **(b)** 1 rps and **(c)** 2 rps.

4. Conclusions

The orientation of microcrystallites has been observed in polycrystalline foil samples of Hf, α-ZrHf and α-TiHf by TDPAC measurements and further confirmed by XRD measurements. It is conjectured that the TDPAC measurements done at an unique positioning of the sample following each step of the annealing treatments, would give an indication about the growth of microcrystallites if any in the sample. The experimental ways of arriving at the polycrystalline analogue of the texturized samples have been discussed. A new method of averaging the $R(t)$ spectra obtained at different positions of the sample to obtain the polycrystalline analogue of the textured sample has been proposed.

Acknowledgements

We thank J Jayapandian for the fabrication of a sophisticated sample rotating set-up for TDPAC measurements and G V N Rao for XRD measurements.

References

[1] E Karlsson, E Mathias and K Siegbahn (eds), *Perturbed angular correlation* (North-Holland, Amsterdam, 1962)
[2] H Frauenfelder and R M Steffan, in *Alpha-, Beta- and Gamma-ray spectroscopy* edited by K Siegbahn (North-Holland, Amsterdam, 1968)
[3] A Weidinger, R Wessner, Th Wichert and E Recknagel, *Phys. Lett.* **A72**, 369 (1979)
[4] M Rots, R Coussement, J Claes and L Hermans, *Hyp. Int.* **11**, 185 (1981)
[5] Yoichi Kawase, Sin-ichi Uehara and Saburo Nasu, *J. Phys. Soc. Jpn.* **61**, 3601 (1992)
[6] R L Rasera, T Butz, A Vasquez, H Ernst, G K Shenoy, B D Dunlap, R C Reno and G Schmidt, *J. Phys.* **F8**, 1579 (1979)
[7] R Govindaraj and K P Gopinathan, *J. Nucl. Mater.* **231**, 141 (1996)
[8] A R Arends, C Hohenemser, F Plieter, H De Waard, L Chow and R M Suter, *Hyp. Int.* **8**, 191 (1980)
[9] E Recknagel, G Schatz and Tu Wichert, in *Hyperfine interactions of radioactive nuclei* edited by J Christiansen (Springer Verlag, Heidelberg, 1983)
[10] G S Collins, S L Shropshire and J Fan, *Hyp. Int.* **62**, 1 (1990)
[11] M Deicher, G Grubel, W Reiner and Th. Wichert, *Hyp. Int.* **15/16**, 437 (1983)
[12] R Govindaraj, K P Gopinathan and B Viswanathan, (to be communicated)
[13] R Vianden, *Hyp. Int.* **15/16**, 1081 (1983)
[14] R Govindaraj, K P Gopinathan and B Viswanathan, (to be communicated)
[15] R Govindaraj and K P Gopinathan (to be communicated)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 5,
November 1997
pp. 547–553

Phonon dispersion in aluminium arsenide and antimonide

PRAFULLA K JHA*, SUJATA RATH and SANKAR P SANYAL
Department of Physics, C. V. Raman Building, University of Bhopal, Bhopal 462 026, India
Email: physics@unibpl.mp.nic.in
* Present address: Institut Jaume Almera, Consell Supiror d'Invastigacions Cientifiques (CSIC), 08028 Barcelona, Spain

MS received 13 May 1997; revised 1 August 1997

Abstract. The phonon dispersion curves for aluminium arsenide and antimonide have been investigated by using a deformation bond approximation model. The results obtained from this model are compared with the experimental values wherever it is available. Since there is no complete experimental phonon dispersion curves for AlAs, we could not compare our calculated results, but the results of AlSb have been compared with the inelastic neutron scattering measurements at 15 K. However, we compare the phonon frequencies of AlAs and AlSb at critical points of the Brillouin zone obtained by our calculations and Raman spectroscopy measurements. This model predicts the phonon modes satisfactorily in all the symmetry directions of the Brillouin zone (BZ). The spectrum has similar features as observed in other III–V compound semiconductors.

Keywords. Semiconductor; phonons; lattice dynamics; neutron scattering.

PACS Nos **63.20; 63.10**

1. Introduction

The lattice dynamical properties of tetrahedrally bonded Al-containing III–V compounds are interesting, since understanding its phonon properties is crucial for technological application. These include heterojunction bipolar transistor, Bragg reflector superlattices, solid state lasers and high electron mobility transistors. Information on phonon dispersion is important in considering the electronic conduction, the non-radiative relaxation process of electrons and so on. However, not much studies have been conducted to understand the role played by phonons in these compounds although the AlAs and AlSb are the partners of the vastly used and studied superlattices containing GaAs/AlAs and GaSb/AlSb which are of great interest [1–7]. In all the Al-containing III–V compounds, the experimental data is only available for AlSb at 15 K alone [8]. Recently Strauch *et al* [8] have determined the phonon dispersion curves in symmetry directions for AlSb by using inelastic neutron scattering at 15 K. They have also performed the model calculation by using bond charge model, where the model parameters have been obtained from a least square fit. In a recent paper, Molinas-Mata *et al* [9] have studied the phonon dispersion curves along [110] and [111] direction of BZ and internal stresses in some III–V compounds including AlAs and AlSb by using the planer bond charge model (PBCM). In this calculation, the PBCM has been fitted to the [100] and [111] inelastic neutron scattering data. Also, in the

case of AlAs, the mass approximation method has been adopted to calculate the phonon dispersion curves (PDC) by considering GaAs as the reference material. With respect to AlAs, one of the typical III–V compound semiconductors, no measurement on phonon dispersion has been performed so far, because it is difficult to grow sufficiently large single crystals of good quality which are free from oxidation. In the early stage of studying the phonon structure in AlAs, several workers have reported phonon frequencies at critical points in Brillouin zone (BZ) by various experimental methods [10, 11, 15, 16]. The phonon dispersion of AlAs along the Δ [12–14] and Λ directions have also been obtained using confined optical phonons in GaAs/AlAs superlattices. In recent studies, Spencer *et al* [17] and Wagner *et al* [18] have measured second order Raman bands in AlAs layers on GaAs and reported the overtone frequencies at the X and L-points. Quite recently, Azuhata *et al* [19] have measured the second order Raman spectra in AlAs layer grown on a GaAs substrate by Raman spectroscopy. They have also calculated the phonon dispersion curves by using the modified version of adiabatic bond charge model but the values of parameters have been obtained from *ab initio* calculations for AlSb by using a similar approach. Thus a realistic lattice dynamical model still seems to be useful for the theoretical studies of phonon properties for Al-containing III–V compound semiconductor to understand the phonons in them.

In the present paper, we report the results of the phonon dispersion curves for AlAs and AlSb by using the formalism of a simplified version of the deformation dipole model [22], called deformation bond approximation (DBA) model [23]. In § 2 the model is described briefly, and the results and discussion in § 3.

2. Model

The deformation dipole model is based on the hypothesis that an overall charge rearrangement due to overlap that produces a net concentration of positive charges at the bonds. During lattice vibration, a new charge distribution is set which, in principle, may be described in terms of a multiple expansion. Kunc *et al* [22] incorporated this idea by considering the terms up to dipole which are assumed to be on the ion site. Fifteen parameters take into account the polarization and deformation of electronic orbitals. These fifteen parameters are very difficult to find due to very limited number of input parameters, which is difficult to employ for the study of lattice vibrations. Kunc *et al* [23] have simplified this model for practical purposes and used for several compound semiconductors. This simplified version of deformation dipole model is known as deformation bond approximation and is quite successful in explaining the phonon properties [24]. The simplification consists: (i) In assumption, reducing the number of independent deformabilities. (ii) In the neglect of "nonlocal electronic polarizabilities". By neglecting the nonlocal polarizabilities, we get a model equivalent to that of Karo and Hardy [25, 26]. This approximation is formally equivalent to the shell model in the case of ionic and semi-ionic crystals [22].

The dynamical matrix corresponding to deformation bond approximation (DBA) model is written as [23]

$$C(q) = C^{\mathrm{sr}}(q) - M^{-1/2}(e + N^{+})(I - Ba)(e + N)M^{-1/2}, \tag{1}$$

	Input parameters			Output parameters	
	AlSb	AlAs		AlSb	AlAs
a	6.1300	5.6600	A	−0.3816	−0.4489
ω_{TO}	9.5700	10.8100	B	−0.3150	−0.3212
ω_{LO}	10.2100	12.1300	C_1	−0.0072	−0.0081
ϵ_∞	15.6800		D_1	−0.0061	−0.0067
C_{11}	0.894	—	F_1	0.0010	0.0013
C_{12}	0.443	—	C_2	0.0037	0.0039
C_{44}	0.416	—	D_2	0.0048	0.0055
e_{14}	−2.04	—	F_2	−0.0056	0.0059
α_1	0.221	—	γ_1	0.5711	0.6401
α_2	10.221	—	γ_2	0.0011	0.0017

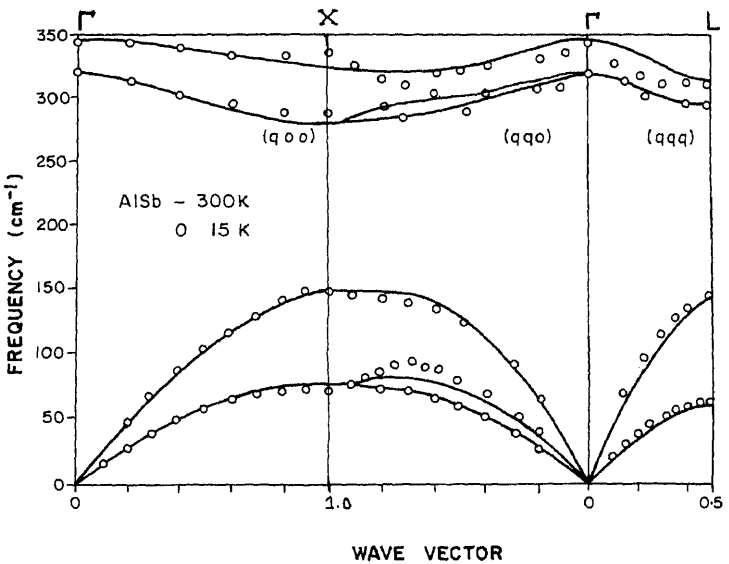


Figure 1. Phonon dispersion curves of AlSb. Experimental points are taken from [8,21].

here I is a unit matrix, $C^{sr}(q)$ is short range matrix, and the matrices N and a are the urier transform of the deformability and polarizability respectively [23, 27]. The matrix in the dynamical matrix is the contribution from the Coulomb coefficients and can be und in [28, 29]. We have also calculated the PDC for II–VI compound semiconductors using the same theoretical approach [30]. The results are in good agreement with periment. The present model has ten parameters and most of them are determined from eir relations with some macroscopic experimental data. These are listed in table 1, here A, B, C_i, D_i, F_i $(i = 1, 2)$ are the short range force constants, and are given in ref. 3]. The input parameters are taken from [31, 32] and presented in table 1 alongwith

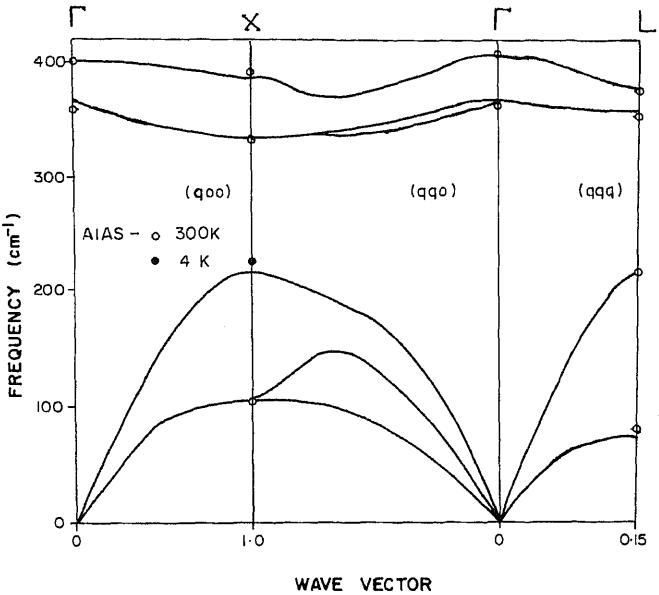


Figure 2. Phonon dispersion curves of AlAs. Experimental points are taken from [19, 34].

output parameters. The notations have their usual meaning. The values of the force constants listed in table 1 are quite reasonable and have not attained any unphysical value though they are partially fitted to macroscopic properties.

3. Results and discussion

The calculated phonon dispersion curves of AlSb by using the model mentioned above are shown in figure 1. The calculated results are compared with the measured data at 15 K [8] due to the non-availability of room temperature experimental data for all the wave vectors in the symmetry directions of Brillouin zone (except at Γ-point). This can also be justified from the fact that the Γ point frequencies are nearly the same at 300 K and 15 K [8]. It is revealed from figure 1 that the calculated results using DBA model agree reasonably well with the available neutron scattering data.

The dispersion of both optical and acoustic phonons are explained more or less satisfactorily in the Δ direction for AlSb. However, minor disagreement for LO and TA modes along X-direction can be observed. Also, the LO and TO branches are almost parallel from Γ to *X* point. The phonon modes in Σ-direction of the Brillouin zone are also reproduced reasonably well from the present model. The other important feature of the PDC, the flattening of the acoustic branches which is observed in almost all the III–V compound semiconductors has been correctly predicted from the present model calculations.

The calculated phonon dispersion curves for AlAs by using the modified version of deformation dipole model are presented in figure 2. For the calculations of PDC of AlSb except the mass and lattice parameters very limited informations about the microscopic

Table 2. Phonon frequencies at critical points of the symmetry direction of the Brillouin zone for AlAs.

Mode	Experiment	Phonon frequency (cm^{-1})		
		Previous work		Present work
		Ab initio calculation[e]	BCM[f]	DBA
LO(Γ)	404[a,g], 404[b,g], 402[f,g]	403	399	400
TO(Γ)	361[a,g], 363[b,g], 360[f,g]	365	363	365
LO(*X*)	400[a], 396[b], 403[c], 391[f]	396	390	385
TO(*X*)	340[a], 338[b], 335[c], 332[f]	338	329	335
LA(*X*)	222[c]	219	215	213
TA(*X*)	103[a], 103[b], 109[c] 104[f]	97	107	104
LO(*L*)	373[d],	375	371	375
TO(*L*)	350[a], 350[b], 349[f]	354	348	352
LA(*L*)	–	214	211	212
TA(*L*)	83[b], 77[f]	74	78	75

[a][17], $T = 36$ K; [b][18], $T = 77$ K; [c][36] $T = 4$ K; [d][18] $T = 300$ K; [e][20]; [f][19], $T = 300$ K; [g]from the first order Raman spectra.

Table 3. Phonon frequencies at critical points of the symmetry direction of the Brillouin zone for AlSb.

Mode	Experiment	Phonon frequency (cm^{-1})		
		Previous work		Present work
		Ab initio calculation[c]	BCM[f]	DBA
LO(Γ)	337[a,e], 352[b], 337[d,e]	334	337	337
TO(Γ)	317[a,e], 318[b], 317[d,e]	316	318	317
LO(*X*)	338[a], 341[b]	342	335	331
TO(*X*)	294[a], 294[b], 290[d]	289	292	288
LA(*X*)	154[b]	152	153	153
TA(*X*)	65[a], 69[b], 70[d]	64	74	74
LO(*L*)	320[b],	325	317	323
TO(*L*)	306[a], 307[b], 309[d]	306	309	307
LA(*L*)	142[a], 147[b]	148	150	144
TA(*L*)	58[a], 56[b], 53[d]	49	55	53

[a][36], $T = 300$ K; [b][8], $T = 15$ K; [c][20] $T = 4$ K; [d][18] $T = 300$ K; [e]from the first order Raman spectra; [f][19], $T = 300$ K.

properties for AlAs are available so far. The phonon dispersion curves for AlAs could also not be compared with the experimental measurements. However, the phonon frequencies at the critical points of the symmetry directions of BZ are listed in table 2 with some experimental data obtained from second order Raman spectra and *ab initio* calculations of Giannozziet *et al* [20]. Thus there is a good agreement between the present calculations and previous experimentally and theoretically obtained values. It can also be seen from figure 2 that all the common features of III–V compound semiconductors have been correctly predicted from the present version of deformation dipole model.

We also present a comparative study of the phonon frequencies for AlAs and AlSb at critical points of the symmetry direction of BZ, calculated by the present model theory and some earlier calculated results in tables 2 and 3 to judge the validity of the present model. It is quite successful in explaining the phonon frequencies at critical points.

Thus we have presented here the results on phonon dispersion curves of AlAs and AlSb using DBA model. The present model has been quite successful in explaining the gross features of PDC of this compound as in the case of other III–V compound semiconductor. We emphasize for the neutron scattering measurements of these compound.

Acknowledgements

The authors are grateful to the Department of Science and Technology and UGC, Govt. of India for financial support. PKJ and SR would like to acknowledge the research associateship and senior research fellowship award respectively from CSIR, New Delhi.

References

[1] B Jusserand, D Paquet and A Regreny, *Phys. Rev.* **B30**, 6245 (1984)
[2] E Molinari, A Fasolino and K Kunc, *Superlatt. and Microstr.* **2**, 397 (1986)
[3] M V Klein, *IEEE J. Quantum Electrton.* **QE22**, 1760 (1986) and references therein
[4] M Cardona, *Superlatt. and Microstr.* **7**, 183 (1990)
M Cardona, *Spectroscopy of semiconductors* (Plenum Press, New York, 1990)
[5] S Rath and S P Sanyal, *Phys. Status Solidi* **B176**, 63 (1993)
[6] G P Schwarz, G S Gualtiers, W A Sunder and L A Farrow, *Superlatt. and Microstr.* **3**, 523 (1987)
[7] P V Santos, A K Sood, M Cardona, K Ploog, Y Ohmori and H Okamoto, *Phys. Rev.* **B37**, 6381 (1988)
[8] D Strauch, B Dorner and K Karch, 'Phonons 89', *Proceeding of Third International Conference on Phonon Physics* (World Scientific, Singapore, 1989)
[9] P Molinas-Mata, A J Shields and M Cardona, *Phys. Rev.* **B47**, 1866 (1994)
[10] M R Lorenz, R Chicotka, D G Pettit and P J Dean, *Solid State Commun.* **8**, 693 (1970)
[11] A Oriton and R Chicotka, *Phys. Rev.* **B10**, 591 (1974)
[12] Z P Wang, D S Jiang and K Ploog, *Solid State Commun.* **65**, 6619 (1988)
[13] A Yu Pusen, A G Mileklin, M P Sinyuleore, K Ploog and A I Tosopore, *JETP* **52**, 462 (1990)
[14] D J Mowbray, M Cardona, K Ploog, *Phys. Rev.* **B43**, 1598 (1991)
[15] Z V Propovic, M Cardona, E Richter, D Strauch, L Taffer and K Ploog, *Phys. Rev.* **B41**, 5904 (1991)
[16] F Calu, P J Mowbray, D W Noles, M Cardona, J M Caluja and K Ploog, *Phys. Rev.* **B45**, 9152 (1991)
[17] G S Spencer, J Grant, R Gray, J Zolman, J Menendez, R Drooped and G N Maracos, *Phys. Rev.* **B49**, 5761 (1994)
[18] J Wagner, A Fisher, W Barus and K Pllog, *Phys. Rev.* **B49**, 7295 (1994)
[19] T Azuhata, T Sota and K Suzuki, *J. Phys.* **C7**, 1949 (1995)
[20] P Giannozi, S de Gironcoli, D Pavone and S Baroni, *Phys. Rev.* **B43**, 7231 (1991)
[21] D Strauch and B Dorner, *J. Phys.* **C19**, 2853 (1986)
[22] K Kunc, M Balakanaski and M Nusimovici, *Phys. Rev.* **B12**, 4346 (1975)
[23] K Kunc, M Balakanaski and M Nusimovici, *Phys. Status Solidi* **B71**, 341 (1975)
[24] K Kunc, M Balakanaski and M Nucimovici, *Phys. Status Solidi* **B73**, 456 (1976)

[25] J R Hardy, *Philos. Mag.* **4**, 1278 (1959); **5**, 859 (1960); **6**, 27 (1961); **7**, 315 (1965)
[26] A M Karo and J R Hardy, *Phys. Rev.* **B129**, 2024 (1965); **141**, 696, (1966); **181**, 1272 (1969)
[27] N Vegelatos, D Wehe and J S King, *J. Chem. Phys.* **60**, 3613 (1974)
R M Martin, *Phys. Rev.* **B1**, 4005 (1970)
[28] J L Yarnell, J L Warren and R G Wenzel, *Neutron scattering* (International Atomic Agency, Vienna, 1968) vol. 1, pp. 301
[29] J F Vetelino and S Mitra, *Phys. Rev.* **B178**, 1349 (1969)
[30] P K Jha, S Rath and S P Sanyal, *Indian J. Pure Appl. Phys.* **34**, 269 (1996)
[31] R W Weikoff, *Crystal structures* (Interscience Publishers, New York, 1963)
[32] E Burstein, H M Brodsky and G Lucovsky, *Int. J. Quantum. Chem.* **15**, 759 (1967)
[33] F S Hickernell and W R Gaytam, *J. Appl. Phys.* **37**, 462 (1966)
[34] G Arlt and P Quadfling, *Phys. Status Solidi* **25**, 323 (1968)
[35] Y S Raptis, E Anastassakis and G Kanellis, *Phys. Rev.* **B46**, 15801 (1992)

-journal of
physics

November 1997
pp. 555–561

lowness surfaces, ray velocity surfaces and phonon ocussing in high T_c superconductor crystals

PHILIP and M S KALA
epartment of Physics and Instrumentation, Cochin University of Science and Technology, ochin 682 022, India

S received 16 June 1997; revised 23 October 1997

bstract. This paper reports the results of the study of anisotropy in elastic wave propagation single crystal superconducting BSCCO. The inverse and group velocities of elastic waves opagating in different directions have been computed and the corresponding slowness and ray locity surfaces plotted, taking elastic constant data from literature. In addition, the phenomenon phonon focussing has been investigated in this material by computing the phonon enhancement ctor along different directions in spherical polar coordinates. The abnormally high values in onon enhancement factor exhibited in certain directions for the phonon modes are interpreted as ie to caustics occurring in the geometrical acoustics approximation adopted in the computational alysis. The results in LSCO and YBCO are found to be similar to those in BSCCO.

eywords. High T_c superconductors; slowness surfaces; ray velocity surfaces; phonon focussing tastrophes.

ACS Nos 74.25; 63.20; 62.20

Introduction

is well known that a detailed description of the nature of elastic wave propaga-on in anisotropic solids can be done by plotting the phase velocity, slowness and y velocity surfaces [1]. Of these, the ray surface is physically the most mean-gful one because it represents the wave front of equal phase for an oscillatory sturbance a unit time interval after it has been created at the origin. Several papers and ooks have appeared in literature depicting the features of these surfaces in different ystal systems and it is found that several crystals exhibit cuspidal edges in ray elocity surfaces along specific directions for selected modes [1–3]. The slowness rface, which is the polar reciprocal of the ray surface, plays an important role in e discussions that follow since one can trace all the important features of ray rfaces from the corresponding slowness surfaces. The analytical techniques for enerating slowness and ray surfaces are well established and the conditions for the istence of cuspidal edges in ray surfaces have been worked out for different crystal asses.

Another important feature of elastic anisotropy is 'phonon focussing', which arises due the non-collinearity of the ray and phase velocities. Whenever the slowness surface is onspherical in shape in any direction, more than one wave vector may correspond to a

single ray velocity vector and the corresponding ray surface will contain folds or cuspidal edges in that direction. Consequently, energy flow in that direction is enhanced compared to other directions. This effect, known as phonon focussing, is now well understood and has been a subject of intense research for the past two decades or so both theoretically and experimentally [4–7].

Several workers have carried out heat pulse experiments in ultrapure single crystals of Ge, GaAs, Si etc. and have demonstrated phonon focussing effects along selected directions for transverse acoustic modes. Most of these experiments exploit temporal analysis of the heat pulse intensity. Results of ballistic phonon imaging experiments carried out on Ge [7], GaAs [8] and KDP [9] show striking differences in the intensity of phonons propagating along different directions which have correspondence with phonon enhancement calculations in these crystals. Ballistic phonon imaging offers a geometric visualization of anisotropic phonon propagation in crystals.

The mechanism of superconductivity and the role played by phonons in superconducting transition in high T_c superconductors remain unclear even today. Of late, with the success in growing single crystals of many of the high T_c superconductors, it has become possible to get a better insight into these materials with the aid of fine experiments. The resonant ultrasound technique has been used to isolate the independent elastic constants of many of these materials [10–12]. However, all the independent elastic constants are available only for a few high T_c superconductor crystals, that too only at room temperature. The only material for which all the elastic constants above and below T_c have been reported is a textured BSCCO (2212) crystal with cylindrical symmetry about the c-axis [13]. Assumption of cylindrical symmetry merges some of the elastic constants resulting in only five independent ones.

In this paper, we trace the slowness and ray surfaces of single crystal BSCCO, taking elastic constant data from literature [13]. Phonon enhancement factors have been evaluated in spherical polar coordinates and plotted. An outline of the computational procedure, results obtained and a discussion of the results are given in the following sections. Even though results have been obtained on YBCO and LSCO, they are not reported here as these are similar to those of BSCCO.

2. Slowness and ray surfaces for BSCCO above and below T_c

The Christoffel equation for the propagation of elastic waves through a crystal is of the form [1]

$$\Omega(\omega, k_x, k_y, k_z) = 0. \tag{1}$$

A vector $\mathbf{m}$, with $\mathbf{m} = \mathbf{n}/v$, having the direction of the wave normal and magnitude equal to the reciprocal of the phase velocity v, is known as the reciprocal velocity vector or slowness vector.

The ray velocity or the velocity with which energy is transported is given by [1]

$$\mathbf{S} = -\frac{\nabla_k \Omega}{(\partial\Omega/\partial\omega)}. \tag{2}$$

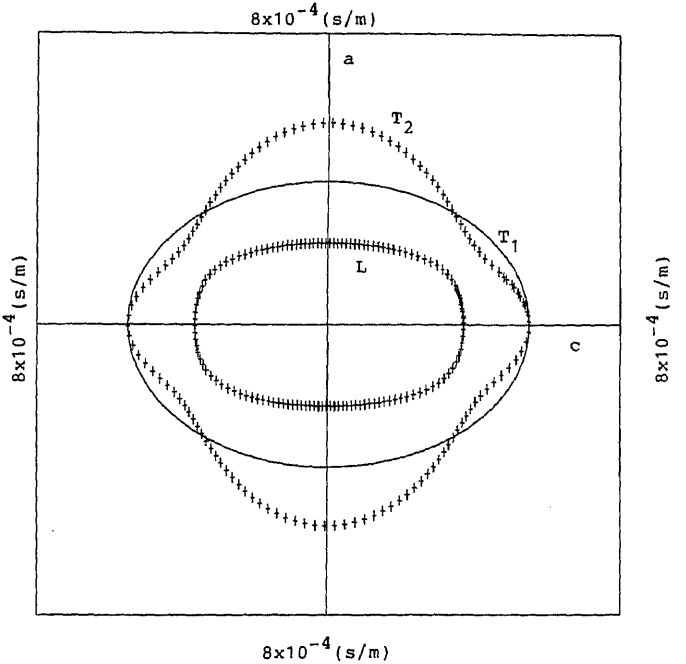


Figure 1. Sections of the slowness surfaces for BSCCO in the a–c plane at 20 K. L, T_1 and T_2 denote the quasilongitudinal, pure shear and quasishear modes respectively. Because of transverse isotropy assumed, the surfaces are concentric circles in the a–b plane. Surfaces for a–c and b–c planes are identical.

Expressions for the components of the slowness and ray velocities for wave propagation in different planes for orthorhombic symmetry have been used to plot the corresponding surfaces for high T_c superconductor crystals. Corresponding to the three solutions of equations (1) and (3), three modes of elastic waves propagate in any direction, which correspond to quasilongitudinal, pure shear and quasishear modes. The slowness and ray velocity surfaces have been plotted for the high T_c superconductor BSCCO in different symmetry planes at two temperatures, one above and the other below T_c, taking elastic constant data from literature [13].

As is expected for an orthorhombic crystal with cylindrical symmetry, wave propagation in the a–b plane is isotropic and the corresponding slowness and ray surfaces are simply concentric circles. The surfaces have been plotted for the a–c plane at 290 and 20 K and those at 20 K are shown in figures 1 and 2. It may be noted that the slowness surface for the quasishear mode deviates from circular shape and contains regions of positive and negative curvatures which give rise to cuspidal edges or folds in the ray surface.

3. Computation of phonon enhancement factor for BSCCO

The direction of energy flux (θ_s, ϕ_s) is obtained from the ray velocity components which in turn depend on the corresponding directions of the phase velocity (θ_k, ϕ_k). This k space

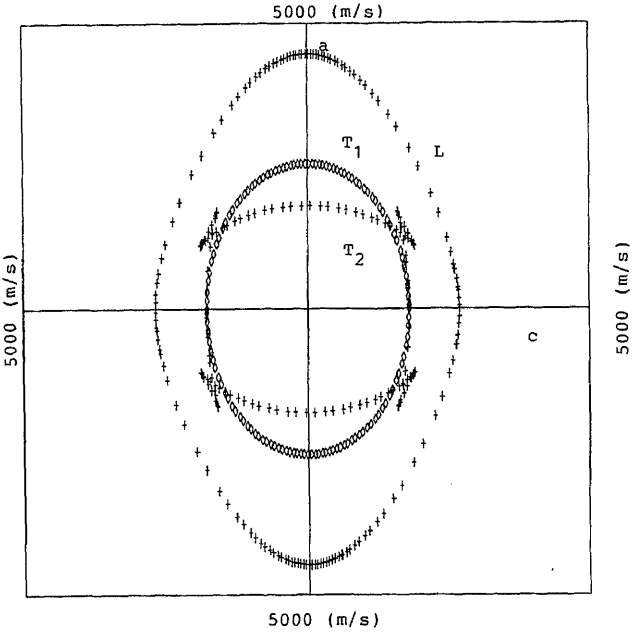


Figure 2. Sections of the ray velocity surfaces for BSCCO at 20 K in the $a-c$ plane. L, T_1 and T_2 have the same meaning as in figure 1.

to s space transformation may be expressed as

$$\begin{aligned} \cos\theta_s &= f(\cos\theta_k, \phi_k), \\ \phi_s &= g(\cos\theta_k, \phi_k), \end{aligned} \tag{3}$$

where the functions f and g are determined from the components of s. Equations (3) are basically a mapping of one two-dimensional space (θ_k, ϕ_k) into another two-dimensional space (θ_s, ϕ_s). The ratio of the product of the differentials in these two spaces is the Jacobian of the functions f and g. This can be written as

$$\mathrm{d}\Omega_s = \mathrm{d}(\cos\theta_s)\mathrm{d}\phi_s = J\mathrm{d}(\cos\theta_k)\mathrm{d}\phi_k = J\mathrm{d}\Omega_k, \tag{4}$$

where the Jacobian is given by

$$J = \begin{vmatrix} \partial f/\partial\cos\theta_k & \partial f/\partial\phi_k \\ \partial_g/\partial\cos\theta_k & \partial_g/\partial\phi_k \end{vmatrix}. \tag{5}$$

This equation provides the desired link between the mathematical formalism and the phonon intensities. The phonon enhancement factor A, defined as the ratio of the solid angle in the phase velocity space to the corresponding solid angle in the ray velocity space, is now related to the Jacobian by the relation [2]

$$A = \left|\frac{\Delta\Omega_k}{\Delta\Omega_s}\right| = \frac{1}{|J|}. \tag{6}$$

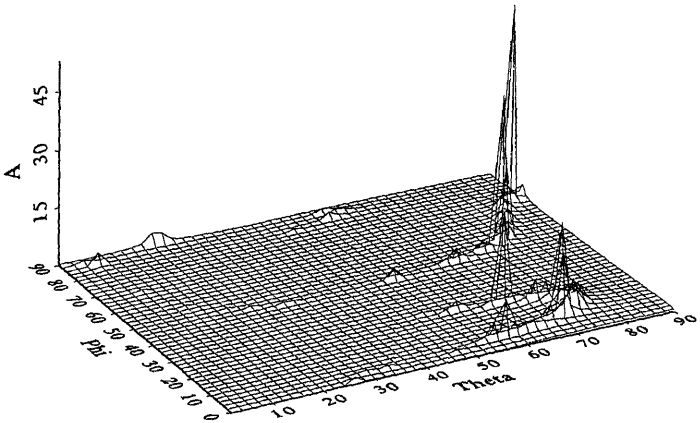


Figure 3. Pseudo-3D representation of the distribution of phonon intensities for the quasishear mode of BSCCO at 20 K. A is the phonon enhancement factor, with θ (theta) and ϕ (phi) designating the polar and azimuthal angles in spherical polar coordinates.

The geometrical interpretation of the Jacobian J is that $J = K|k^2|\cos\zeta$, where K is the Gaussian curvature of the slowness surface and ζ is the angle between the phase velocity vector and ray velocity vector. The Gaussian curvature is the product of the two extremal curvatures (inverse radii) of an elemental surface. If either or both of the extremal curvatures vanish, J and K will vanish. The Jacobian is dimensionless and is directly related to the phonon enhancement factor A by (6).

We have computed the phonon enhancement factors for single crystal BSCCO at 290 and 20 K following the method outlined above. Computations have been performed varying θ_k and ϕ_k from 0 to $\pi/2$, each in steps of 2° resulting in performing the computation at 2025 points in the wave vector space. Figure 3 gives a pseudo-3D view of the phonon intensities distributed in the first quadrant of the wave vector space for the quasishear mode in BSCCO at 20 K. Even though we have done the computation for the pure shear and quasilongitudinal modes as well, these are not reproduced here as no anomalous features are expected to be exhibited by these modes in phonon intensity distribution. However, similar to the peaks observed in the quasishear mode, very sharp peaks are seen in the enhancement factors for the pure shear and quasilongitudinal modes as well, but in other directions. Mirror reflections of this figure give a complete pseudo-3D representation of the phonon intensities distributed in the wave vector space for quasishear mode.

Phonon enhancement factor is found to take physically unrealistic values for all the modes of wave propagation in certain directions. In certain cases the values run into several thousands, particularly for the pure shear mode. This can be attributed only to singularities occurring in the phonon enhancement factor. As has been described by several workers before [14, 15], these singularities are a consequence of the geometrical acoustics approximation used in the definition and evaluation of phonon enhancement factor. Note that the ray surfaces for the pure shear mode do not contain folds or cusps and so one cannot expect any anomalous values for the phonon enhancement factor for

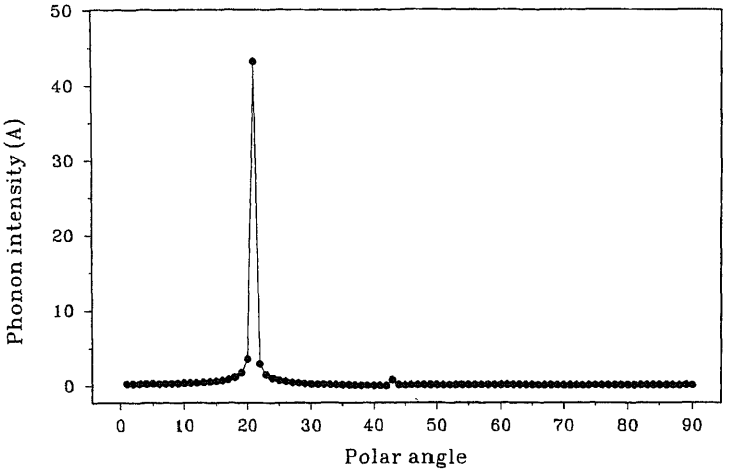


Figure 4. Variation of phonon intensity with the polar angle θ for the propagation of quasishear mode in the a–c plane of BSCCO at 20 K. The peak in the figure corresponds to the point at which caustics intersect the a–c plane.

this mode. This is true for the quasilongitudinal mode as well. The only mode for which we can expect to see interesting features in phonon enhancement is the quasishear mode.

We have computed the phonon enhancement factors along the cusp directions for the quasishear mode. The general program developed to compute phonon enhancement factors along general directions has been modified to compute phonon enhancement factors along the directions in which cusps occur in the ray velocity surface. The computation have been performed on BSCCO and the 20 K result is shown in figure 4. The directions in which peaks in phonon intensities occur in this figure correspond to the points at which the caustics intersect the (101) plane in this crystal. A qualitative analysis of the positions of these peaks in comparison with the directions in which cusps occur indicate that these are the directions along which the phonon intensity is getting amplified. The values of phonon enhancement factors are not abnormally high, and these are the directions in which one should look for phonon focussing effects in an experiment. Since the results in YBCO and LSCO are similar to these, they are not reproduced here.

4. Conclusions

We have computed and plotted the slowness surfaces, ray surfaces and phonon enhancement factors in single crystals of the high T_c superconductors BSCCO, YBCO and LSCO. The results bring out the changes that the slowness and ray surfaces undergo when the material undergoes superconducting transition. The results indicate that there are no significant changes in these surfaces due to superconducting transition or superconducting transition does not significantly affect the elastic anisotropy of these materials.

It is known that phonon focussing does not have any direct bearing on superconducting transition. All we see is a change in the magnitude of the phonon enhancement factor at temperatures above and below T_c. This change merely reflects the changes that the elastic

constants undergo over this temperature range. We have identified the directions in which folds occur in the ray velocity surfaces and verified that phonon intensity for the quasishear mode gets enhanced in these directions. The phonon mean free path near T_c is very small in these materials and one cannot expect to see phonon focussing effects at such elevated temperatures. One must look for these effects in ultrapure single crystals at ultralow temperatures, or under the boundary scattering regime.

References

[1] B A Auld, *Acoustic fields and waves in solids* (John Wiley and Sons, New York, 1973) vol. 1
[2] J Philip and K S Viswanathan, *Pramana – J. Phys.* **8**, 348 (1977)
[3] A G Every, *Phys. Rev.* **B24**, 3456 (1981)
[4] B Taylor, H J Maris and C Elbaum, *Phys. Rev.* **B3**, 1462 (1971)
[5] C G Winterheimer and A K Mc Curdy, *Phys. Rev.* **B18**, 6576 (1978)
[6] J Philip and K S Viswanathan, *Phys. Rev.* **B17**, 4969 (1978)
[7] G A Northorp and J P Wolfe, *Phys. Rev.* **B22**, 6196 (1980)
[8] R G Ulbrich, V Narayanamurti and M A Chin, *Phys. Rev. Lett.* **45**, 1432 (1980)
[9] M A Weilert, M E Msall, A C Anderson and J P Wolfe, *Phys. Rev. Lett.* **71**, 735 (1993)
[10] J D Maynard and M J Mckenna, *Physical acoustics* edited by M Levy (Academic Press Inc., San Diego, 1992) vol. XX
[11] A Migliori, W M Visscher, S Wong, S E Brown, I Tanaka, H Kojima and P B Allen, *Phys. Rev. Lett.* **64**, 2458 (1990)
[12] Ming Lei, J L Sarrao, W M Visscher, T M Bell, J D Thompson and A Migliori, *Phys. Rev.* **B47**, 6154 (1993)
[13] G A Saunders, C Fanggao, L Jiaqiang, Q Wang, M Cankurtaran, E F Lambson, P J Ford and D P Almond, *Phys. Rev.* **B49**, 9862 (1994)
[14] P Taborek and D L Goodstein, *Solid State Commun.* **33**, 1191 (1980)
[15] S Tamura, *Phys. Rev.* **B28**, 897 (1983)

RAMANA
-journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 563–580

Darboux transformation and elementary exact solutions of the Schrödinger equation

LADISLAV G BAGROV* and BORIS F SAMSONOV
omsk State University, 36 Lenin Ave, 634 050 Tomsk, Russia
High Current Electronics Institute, 4 Akademichesky Ave, 634 055 Tomsk, Russia

IS received 28 February 1996; revised 4 June 1996

bstract. Darboux transformation is applied to three classical potentials, namely the harmonic scillator, effective Coulomb and Morse potentials to generate exactly solvable potentials of ementary form. For every potential, the isospectral families of potentials are constructed. For most all potentials, a set of normalized discrete spectrum wave functions is given.

eywords. Schrödinger equation; Hamiltonian; wavefunction; Darboux transformation.

ACS No. 03.65

Introduction

nce the first reported work [1], the Darboux transformation has been an area of great terest. Perhaps this interest was first reflected in [2], then this method was rediscovered the factorization method [3,4]. The next stimulus to its development was given by the persymmetric formulation of the Schrödinger equation [5]. (For a recent review in this omain see [6]). The soliton theory must be mentioned especially because the Darboux ansformation played a considerable role in the development of this theory (see [7] and ferences therein). This transformation was investigated recently in connection with the ectral theory of the Schrödinger operator by the techniques of dressing chain [8] and as a urce of new exactly solvable potentials [9–11]. Despite the fact that this type of trans-rmation has been studied for a long period, the number of quantum mechanical problems xactly solved in elementary functions is not too large, especially if the question consists finding the normalized discrete spectrum wave functions. In this article we make an tempt to fill this gap. We give and investigate new exactly solvable potentials for the one-mensional Schrödinger equation as applied to the harmonic oscillator, effective Coulomb nd Morse potentials. On the base of each potential we construct a family of isospectral otentials. For almost all potentials we give a set of normalized discrete spectrum wave nctions. We concentrate our attention on the solutions of elementary form.

Generating properties of the Darboux transformation

this section we give a brief overview of the main properties of the Darboux ansformation discussed in a number of earlier publications.

Let us consider the one-dimensional Schrödinger equation (setting $\hbar = 2m = 1$)

$$H_0\psi_E(x) = E\psi_E(x), \quad H_0 = -\mathrm{d}^2/\mathrm{d}x^2 + V_0(x), \quad x \in [a, b]. \tag{1}$$

The potential $V_0(x)$ is supposed to be a real and sufficiently smooth function in the interval $R = [a, b]$ which can be infinite. It is also supposed that we know the general solution of this equation for all values of the parameter E (complex in general). Let functions $u_i(x)$ called transformation functions be particular solutions of (1) with α_i as eigenvalues. Using these functions we can introduce the following operator called the N-order Darboux transformation operator.

$$L_N = W^{-1}(u_1, u_2, \ldots, u_N)\begin{vmatrix} u_1 & u_2 & \cdots & 1 \\ u_1' & u_2' & \cdots & \mathrm{d}/\mathrm{d}x \\ \vdots & \vdots & \ddots & \vdots \\ u_1^{(N)} & u_2^{(N)} & \cdots & \mathrm{d}^N/\mathrm{d}x^N \end{vmatrix}, \tag{2}$$

where $W(u_1, u_2, \ldots, u_N)$ is the Wronskian of functions $u_1, u_2, \ldots, u_N$, $(W(u_1) = u_1)$, the prime denotes a derivative with respect to x, and the determinant is a differential operator obtained by the development of the determinant in the last column with the functional coefficients placed before the derivative operators.

Operator L_N has the property [11]

$$L_N H_0 = H_N L_N, \tag{3}$$

where

$$H_N = -\mathrm{d}^2/\mathrm{d}x^2 + V_N(x) \tag{4}$$

is a new Hamiltonian and the potential difference $A_N(x) = V_N(x) - V_0(x)$ is given by the formula [12, 13]:

$$A_N(x) = -2\frac{\mathrm{d}^2}{\mathrm{d}x^2}\log W(u_1, \ldots, u_N). \tag{5}$$

The operator L_N when applied to any solution of (1) gives the known Crum–Krein formula [12, 13] for the solutions of a new Schrödinger equation with the Hamiltonian H_N

$$\varphi_E(x) = L_N\psi_E(x). \tag{6}$$

It can be shown [8, 11] that any N-order differential transformation operator satisfying (3) can always be presented in the form (2) by properly choosing its arbitrary constants.

If H_0 and H_N are formally self-adjoint operators (in the sense of some scalar product) then we have from formula (3)

$$H_0 L_N^+ = L_N^+ H_N. \tag{7}$$

It follows from this relation that the operator $L_N^+ \equiv L_N^\dagger$ [being Hermitian conjugated to L_N with respect to an inner product for which $(\mathrm{d}/\mathrm{d}x)^+ = -(\mathrm{d}/\mathrm{d}x)$] realizes the transformation in the inverse direction i.e. from the solutions of the new Schrödinger equation to the solutions of the initial one, and consequently the product $L_N^+ L_N$ is a symmetry operator for (1). In full analogy, the operator $L_N L_N^+$ is a symmetry operator for the new Schrödinger equation. Since we deal with a one dimensional Schrödinger

equation the above-mentioned symmetry operators are the Hamiltonian and its polynomial functions. As a result, we have the following remarkable factorization properties [14, 11]:

$$L_N^+ L_N = (H_0 - \alpha_1)(H_0 - \alpha_2)\cdots(H_0 - \alpha_N), \tag{8}$$

$$L_N L_N^+ = (H_N - \alpha_1)(H_N - \alpha_2)\cdots(H_N - \alpha_N). \tag{9}$$

In the case $N = 1$ we have the conventional Darboux transformation [1]

$$L \equiv L_1 = -u'(x)/u(x) + \mathrm{d}/\mathrm{d}x, \quad H_0 u = \alpha u$$

which is the first order transformation. The relations (8) and (9) take for this case the form $L^+L = H_0 - \alpha$, $LL^+ = H_1 - \alpha$. These are the properties that are used in the supersymmetric quantum mechanics when one defines the superhamiltonian

$$H = \begin{pmatrix} H_0 & 0 \\ 0 & H_1 \end{pmatrix}$$

and supercharge operators

$$Q = \begin{pmatrix} 0 & 0 \\ L & 0 \end{pmatrix}, \quad Q^+ = \begin{pmatrix} 0 & L^+ \\ 0 & 0 \end{pmatrix}$$

which form superalgebra $sl(1/1)$ [5] with commutation relations $[H, Q] = [H, Q^+] = 0$ and anticommutation ones $\{Q, Q\} = \{Q^+, Q^+\} = 0, \{Q, Q^+\} = H - \alpha I$, where I is the unit 2×2 matrix. Note that the use of the N-order transformation operator L_N and its conjugate L_N^+ in supercharges Q and Q^+ gives N-order superalgebra [11, 14]

$$\{Q, Q^+\} = (H - I\alpha_1)(H - I\alpha_2)\cdots(H - I\alpha_N),$$

where superhamiltonian H has to be constructed with the help of the Hamiltonians H_0 and H_N and I is $(N + 1) \times (N + 1)$-dimensional unit matrix.

It follows from the definition of the operator L that Ker $L = \mathrm{span}\{u\}$ and Ker $L^+ = \mathrm{span}\{\nu\}$ where $\nu = u^{-1}$ and span denotes the linear hull over the complex number field. Note that the function ν is an eigenfunction of the Hamiltonian H_1 and it corresponds to the eigenvalue α. Let $\tilde{u}$ and $\tilde{\nu}$ be another linearly independent solutions of the Schrödinger equations with the Hamiltonians H_0 and H_1 respectively corresponding to the same eigenvalue α. If they are chosen such that $W(u_1, u_2) = 1$ and $W(\nu_1, \nu_2) = 1$ then

$$\tilde{u} = u \int u^{-2}\mathrm{d}x, \quad \tilde{\nu} = \nu \int \nu^{-2}\mathrm{d}x$$

and $L\tilde{u} = \nu, L^+\tilde{\nu} = -u$. This signifies that with the help of the operators L and L^+ we can establish the one-to-one correspondence between the solutions of the initial Schrödinger equation with the Hamiltonian H_0 and the transformed one with the Hamiltonian H_1. Every two dimensional space T_E^0 of the solutions of the input equation with the eigenvalue $E \neq \alpha$ is transformed in the space T_E^1 of the solutions of the final equation with the same eigenvalue E. If $E = \alpha$ we have $\tilde{u} \to \nu, \tilde{\nu} \to -u$. This means that the Darboux transformation operator L gives a complete information on the transformed Schrödinger equation. In particular we can obtain not only the discrete spectrum eigenfunctions but the scattering data for the Hamiltonian H_1 as well.

The scattering data for the Schrödinger equation (1) considered on full axis $R = (-\infty, +\infty)$ can be obtained with the help of its Jost solutions $\psi^+(x,k)$ and $\psi^-(x,k)$ (see for example [15]) with $k^2 = -E$ and k is a complex number. These solutions have the following asymptotic behavior $\psi^{\pm}(x,k) \to \exp(\mp kx)$ as $x \to \pm\infty$ at $\mathrm{Re}(k) > 0$. We suppose that the discrete spectrum of (1) (if exists) is negative ($E_i < 0$) and the continuous one is positive. The four Jost solutions $\psi^{\pm}(x,k)$ and $\psi^{\pm}(x,-k)$ are linearly dependent from each other, namely any of them can be expressed as a linear combination of any two others. One can use $\psi^+(x,k)$ and $\psi^+(x,-k)$ as the basis in the two-dimensional space of the solutions of (1) at $E = -k^2$ since their Wronskian equals $2k \neq 0$. Then $\psi^-(x,k)$ can be expanded in terms of this basis as follows [9]:

$$\psi^-(x,k) = a_0(k)\psi^+(x,-k) + b_0(k)\psi^+(x,k), \quad E = -k^2, \quad \mathrm{Re}\, k > 0.$$

Coefficients $a_0(k)$ and $b_0(k)$ define the scattering data [15] for the potential $V_0(x)$. For the real valued potentials these coefficients satisfy the following conditions

$$a_0(-k) = a_0^*(k), \quad b_0(-k) = b_0^*(k), \quad |a_0(k)|^2 - |b_0(k)|^2 = 1,$$

where the asterisk implies the complex conjugation.

Consider first the case when the transformation function u is the ground state function of the Hamiltonian H_0: $H_0 u = \alpha u$, $E_0 = \alpha = -k_0^2 < 0$. In this case the energy E_0 does not belong to the discrete spectrum of the transformed Hamiltonian. Using the known asymptotic behavior of the function u we find the scattering data for the Hamiltonian H_1. (Note that differentiating of the asymptotic of functions can be completely justified in this case). Since $L^+L = H_0 - \alpha$ and $LL^+ = H_1 - \alpha$ the Jost solutions for the two Schrödinger equations with the Hamiltonian H_0 and H_1 are interrelated by the operators L and L^+ [9]

$$\varphi^{\pm}(x,k) = \pm(k_0 - k)^{-1} L\psi^{\pm}(x,k), \quad \psi^{\pm}(x,k) = \pm(k_0 + k)^{-1} L^+\psi^{\pm}(x,k),$$
$$k_0^2 = -\alpha > 0.$$

It follows from these relations that the coefficients $a_1(k)$ and $b_1(k)$ defining the scattering data for the Hamiltonian H_1 are expressed in terms of the coefficients $a_0(k)$ and $b_0(k)$ as follows [9]:

$$a_1(k) = \frac{k + k_0}{k - k_0} a_0(k), \quad b_1(k) = -b_0(k).$$

If we insert a new eigenvalue in the spectrum of the initial Hamiltonian we should replace $k_0 \to -k_0$ since this procedure is equivalent to deletion of the ground state level from the Hamiltonian H_1.

It can be shown [11] that the N-order operator L_N can always be presented as a product of N first order Darboux transformation operators: $L_N = L^{(N-1,N)} L^{(N-2,N-1)} \cdots L^{(0,1)}$ and consequently the formula (2) represents the final result of the iteration of N first order transformations. Every operator $L^{(p-1,p)}$ intertwines the Hamiltonians H_{p-1} and H_p: $L^{(p-1,p)} H_{p-1} = H_p L^{(p-1,p)}, p = 1, 2, \ldots, N$ and indeed we have a chain of Hamiltonians $H_0 \to H_1 \to \cdots \to H_N$ obtained one from another by the chain of sequential transformations $L^{(0,1)}$, $L^{(1,2)}$, $L^{(N-1,N)}$ (so called dressing chain, see for example [8]). Some properties of such a chain was recently studied in connection with the construction of reflectionless potentials with infinite discrete spectrum [9].

In the two cases the chain so obtained can have ill-defined elements [16] although the final Hamiltonian remains well-defined. The first case which corresponds to complex-valued intermediate potentials has been described in [14]. We can illustrate this case by the following simple example [17]. It can be shown that for $N = 2$ the formula (5) gives the real-valued potential if the (essentially complex) transformation functions u_1 and u_2 correspond to mutually conjugated eigenvalues: $H_0u_1 = \alpha u_1, H_0u_2 = \alpha^* u_2$. The intermediate transformations give in this case the complex-valued potential differences and the intermediate Hamiltonian H_1 is not self-adjoint.

In the second case we can obtain using (5) a regular potential $V_2(x) = V_0(x) + A_2(x)$ using two juxtaposed discrete energy spectrum eigenfunctions $u_1(x) = \psi_j(x)$ and $u_2(x) = \psi_{j+1}(x)$ of the Hamiltonian H_0 as transformation functions. This fact was proved by Krein [13] and recently rediscovered by Adler [18]. The intermediate potential $V_1(x)$ is in this case a singular function in the interval R and the spectral problem for it strongly differs from the one for the Hamiltonians H_0 and H_2. Using transformations of these kind we can construct a supersymmetric quantum mechanical model with unusual properties [19], namely, the states which are annihilated by supercharges are located in the middle of the discrete spectrum of the superhamiltonian and the ground state of the superhamiltonian is degenerate.

In the general case of N transformation functions $u_{m_1}, u_{m_2}, \ldots, u_{m_N}$ the potential difference [see the formula (5)] is a regular function if the Wronskian $W(u_{m_1}, u_{m_2}, \ldots, u_{m_N})$ conserves its sign on the interval R. For the case when $u_{m_i}, i = 1, 2, \ldots, N$, are the discrete spectrum eigenfunctions of the Hamiltonian H_0 the sign conservation condition of this Wronskian is known [13]. The Wronskian $W(u_{m_1}, u_{m_2}, \ldots, u_{m_N})$ conserves its sign for $u_{m_i} \in L^2$ (the space of square integrable functions on the interval R) if the integers $(0) \leq m_1 < m_2 < \cdots < m_N$, being equal to a number of zeros of the function u_{m_i}, satisfy the condition $(m - m_1)(m - m_2) \cdots (m - m_N) \geq 0$ for all $m = 0, 1, 2, \ldots$. This condition is satisfied, in particular, if α_{m_i} are two by two, the juxtaposed points of the discrete spectrum. Using this property we can construct exactly solvable potentials in terms of elementary functions. The levels with $E = \alpha_{m_i}$ will be absent in the discrete spectrum of the transformed Hamiltonians. Nevertheless the functions $\varphi_i = L_N\psi_i$ form the discrete basis of the space L^2 [13, 18].

It follows from the formulae (6) and (2) that $L_Nu_i = 0, i = 1, 2, \ldots, N$ i.e. $\mathrm{Ker}\, L_N = \mathrm{span}\{u_i, i = 1, 2, \ldots, N\}$. Nevertheless, with every solution u_i of the initial Schrödinger equation we can associate the function

$$\varphi_m = W^{(m)}(u_1, u_2, \ldots, u_N)W^{-1}(u_1, u_2, \ldots, u_N) \tag{10}$$

which is the solution of the Schrödinger equation with the potential $V_N(x) = V_0(x) + A_N(x), A_N(x)$ being calculated using the formula (5). $W^{(m)}(u_1, u_2, \ldots, u_N)$ is the Wronskian of the functions $u_1, u_2, \ldots, u_N$ except the function u_m. Since the operator L_N^+ assures the transformation in the inverse direction we have $L_N^+\varphi_m = 0$, $m = 1, 2, \ldots, N$, i.e. $\mathrm{Ker}\, L_N^+ = \mathrm{span}\{\varphi_m, m = 1, 2, \ldots, N\}$. If the functions φ_m satisfy the boundary condition of the discrete spectrum eigenfunctions of the Hamiltonian H_N operator L_N creates N additional discrete spectrum levels for this Hamiltonian with respect to the Hamiltonian H_0.

Operator L_N completely defines the properties of the Hamiltonian H_N. For example, using the factorization property (8) we can easily obtain the normalization constants for

$$\langle\varphi_i|\varphi_i\rangle = \langle L_N\psi_i|L_N\psi_i\rangle = \langle\psi_i|L_N^+L_N\psi_i\rangle = \prod_{m=1}^{N}(E_i - \alpha_m). \tag{11}$$

The same formula permits one to find the coefficients $a_N(k)$ and $b_N(k)$ for the Hamiltonian H_N. For example, if all the transformation functions appertain to the discrete spectrum of the initial Hamiltonian we obtain

$$a_N(k) = \prod_{m=1}^{N}\frac{k+k_m}{k-k_m}\,a_0(k), \quad b_N(k) = (-1)^N b_0(k), \quad k_m^2 = -\alpha_m > 0.$$

Let us now consider the scattering data for the radial Schrödinger equation for the Coulomb potential. One dimensional spectral problem for this case should be considered on the half-line $R = [0,\infty)$ and the potential has an additional term $l(l+1)x^{-2}$. The notation for k used in the papers [20, 21] is different from the one introduced here. To facilitate the comparison of our results with that of [21], we change now the notation for $k : k^2 = E > 0$ for the continuous spectrum of the radial Schrödinger equation. The presence of the centripetal member $l(l+1)x^{-2}$ makes the analysis slightly different from the one on full-line. Two linearly independent Jost solutions depending now on the angular quantum number l, $f_l(k,r)$ and $f_l(-k,r)$ are defined by the following asymptotic behavior at $r \to \infty$ [20]

$$f_l(k,r) \to \mathrm{e}^{(il\pi)/2}\mathrm{e}^{ikr}.$$

The regular at $r = 0$ solution of the Schrödinger equation defined by the boundary condition

$$\lim_{r\to\infty}(2l+1)!!\,r^{-l-1}\psi_l(k,r) = 1$$

is decomposed in terms of the basis functions $f_l(k,r)$ and $f_l(-k,r)$ with the help of the Jost function $F_l(k)$ [20]

$$\psi_l(k,r) = \tfrac{1}{2}ik^{-l-1}[F_l(k)f_l(-k,r) - (-1)^l F_l(-k)f_l(k,r)]. \tag{12}$$

Let us choose the ground state function of the initial Schrödinger equation as the transformation function: $u = \psi_0(k,r), H_0u = \alpha u,\ k_0^2 = -\alpha = -E_0 > 0$. The transformed equation differs in this case from the initial one only by the value of l which has to be changed as follows $l \to l+1$. Using the known asymptotic behavior of the function u and the factorization property $L^+L = H_0 - \alpha$ we find

$$Lf_l^0(k,r) = (k - ik_0)f_{l+1}^1(k,r), \quad L^+f_{l+1}^1(k,r) = (k+ik_0)f_l^0(k,r), \tag{13}$$

where the value of superscript is related with the number of Hamiltonian. Moreover, for the case under consideration the Jost solutions for two Hamiltonians are related as follows: $f_{l+1}^1 = f_{l+1}^0 = if_l^0$. Using the asymptotic behavior of the regular solution $\psi_l^0(k,r)$ at $k \to \infty$ [20]:

$$\psi_l^0(k,r) \to k^{-l-1}\sin(kr - \tfrac{1}{2}l\pi)$$

it is not difficult to obtain the following relations:

$$L\psi_l^0(k,r) = (k^2 + k_0^2)\psi_{l+1}^1(k,r), \quad L^+\psi_{l+1}^1(k,r) = \psi_l^0(k,r). \tag{14}$$

It follows from the formulae (12), (13) and (14) that

$$F_{l+1}^1(k) = -\frac{k}{k - ik_0}F_l^0(k). \tag{15}$$

This formula differs from that given in [21] and obtained with the help of the integral representation of the Jost function only by the sign. This difference is due to the different definition of the tranformation operator: $L = -A^+$. The extension of the formula (15) to the iteractive procedure is obvious. Using the asymptotic behavior of other transformation functions we can find the Jost function for other transformations.

We will now make some comments on the possibility of getting multiparameter families of isospectral potentials. Such a possibility has earlier been discussed in a number of publications, for example [22–27] (see as well [6]). In [28], the author described the possibility of using the differential transformation operators to obtain a one parameter family of potentials which are identical to the ones obtained using the Gelfand–Levitan–Marchenko integral operator. The paper [16] contains the generalization of this approach making possible to get N-parameter family of potentials. As an illustration of the method a new representation of the known N-soliton potential (e.g. N-parametric potential with N arbitrary disposed discrete energy spectrum levels) has been obtained [16]. In particular case $N = 2$ this potential completely coincides with the one reported in [26].

The possibility of getting the family of isospectral potentials resides on the use of the general solution of the initial Schrödinger equation which is given $E = \alpha < E_0$ a linear combination of its two linearly independent solutions as the transformation function u. The position of the discrete energy levels does not depend on the value of the coefficients of this linear combination. Hence, we obtain strictly isospectral potentials having the same discrete spectrum levels and reflection coefficients. If the function $\nu = u^{-1}$ is square integrable (this is the ground state function for the whole family in this case) the potentials differ only by the value of the normalization coefficient of the function ν. Its calculation method is described in [16].

3. Elementary exactly solvable potentials

If the initial Schrödinger equation has elementary solutions, the double Darboux transformation [see formulae (2), (5) and (6) at $N = 2$] with the adjacent points of the discrete spectrum functions as transformation functions gives a new elementary exactly solvable potential. Almost all elementary solutions in one-dimensional quantum mechanics are related with the hypergeometric function or its confluent form. The general potential exactly solvable in hypergeometric functions is cited in [29]. The application of the double Darboux transformation to this potential leads to very complicated formulae. We shall consider now some special simpler cases corresponding to the harmonic oscillator, effective Coulomb and Morse potentials.

3.1 *Harmonic oscillator*

Let H_0 be given by $H_0 = -\mathrm{d}^2/\mathrm{d}x^2 + x^2/4 - 1/2$. Consider first a case where the transformation function does not belong to the space L^2. There exist a family of elementary solutions of the Schrödinger equation for the harmonic oscillator. For example, for even $m = 2p$, we can choose

$$u_m(x) = \psi_{-m}(x) = \exp(x^2/4)H_m(ix/\sqrt{2}), \quad H_0u_m(x) = -(m+1)u_m(x)$$
$$m = 0, 1, 2, \ldots, \qquad (16)$$

where $H_m(x)$ are the Hermite polynomials [30]. The Darboux transformation with these functions produces the family of potentials

$$V_1^{(2p)}(x) = x^2/4 - 3/2 + 8p^2[q_{2p-1}(x)/q_{2p}(x)]^2 - 4p(2p-1)q_{2p-2}(x)/q_{2p}(x),$$
$$q_p(x) = (-i)^p He_p(ix), \quad He_p(x) = 2^{-p/2}H_p(x/\sqrt{2}), \qquad (17)$$
$$q_{p+1}(x) = xq_p(x) + pq_{p-1}(x), \quad q_0(x) = 1, \quad q_1(x) = x.$$

Potential (17) has been studied in detail for $m = 2$ by Dubov *et al* [31]. These potentials have been previously obtained by us [11] but their eigenfunctions were not cited. All these potentials have a sharp minimum in the region $x = 0$ and with growing x they tend to the parabola $x^2/4$. The discrete spectrum of the new Hamiltonians $H_1^{(m)} = -\mathrm{d}^2/\mathrm{d}x^2 + V_1^{(m)}(x)$ has one additional level $E_0 = -(m+1)$ with respect to the Hamiltonian H_0. The discrete spectrum wave functions have the form

$$\varphi_0(x) = \sqrt{(2p)!}(2\pi)^{-1/4}\exp(-x^2/4)/q_{2p}(x),$$
$$\varphi_{n+1}(x) = (2\pi)^{-1/4}(n!)^{-1/2}(n+2p+1)^{-1/2}\exp(-x^2/4) \qquad (18)$$
$$\times [He_n(x)q_{2p+1}(x)/q_{2p}(x) - nHe_{n-1}(x)], \quad n = 0, 1, 2, \ldots.$$

With odd $m = 2p + 1$, functions (16) become zero only for $x = 0$, and one can use them as transformation functions in the region $x \geq 0$. This yields the family of the new potentials:

$$V_1^{(2p+1)}(x) = x^2/4 - 3/2 + 2(2p+1)^2[q_{2p}(x)/q_{2p+1}(x)]^2$$
$$- 4p(2p+1)q_{2p-1}(x)/q_{2p+1}(x).$$

As a solution of the new Schrödinger equation for $E = n$ $(n = 0, 1, 2, \ldots)$, we obtain the functions

$$\varphi_n(x) = \sqrt{2}(2\pi)^{-1/4}(n!)^{-1/2}(n+2p+2)^{-1/2}$$
$$\times [He_n(x)q_{2p+2}(x)/q_{2p+1}(x) - nHe_{n-1}(x)]\exp(-x^2/4). \qquad (19)$$

When n is odd, these functions are the solutions of the new Schrödinger equation becoming zero for $x = 0, \infty$, and when n is even, they are singular for $x = 0$. Due to this fact, the discrete spectrum has only odd n values. Functions (19) are normalized-at-unity on the interval $(0, \infty)$. When $n = 1$, we have the ground state function.

The second order Darboux transformation (see the formulae (2), (5) and (6) at $N = 2$) with the discrete spectrum functions $\psi_p(x) = \exp(-x^2/4)He_p(x)$ and $\psi_{p+1}(x)$ produces

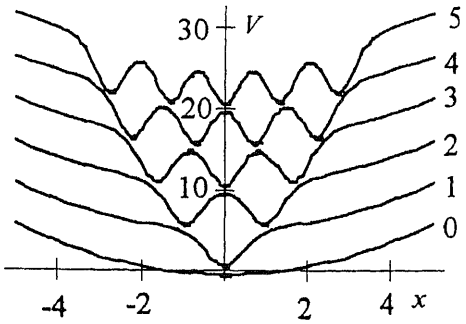


Figure 1. The k-well oscillator-like potentials.

ıer family of potentials:

$$V_2^{(p,p+1)}(x)=\frac{x^2}{4}+\frac{3}{2}-2\frac{J_p''(x)}{J_p(x)}+2\left(\frac{J_p'(x)}{J_p(x)}\right)^2, \quad J_p(x)=\sum_{i=0}^{p}\frac{\Gamma(p+1)}{\Gamma(i+1)}He_i^2(x).$$

e potentials are even functions of x and resemble the parabola $x^2/4$ with N shallow na of equal height at the bottom. We plot in figure 1 the functions $V=V_2^{(p,p+1)}+5p$ e p is equal to the curve number. Their discrete spectrum differs from that of the onic oscillator potential by the absence of the levels $E=p$ and $E=p+1$.

e set of normalized-at-unity wave functions for these potentials can easily be ıed from formulae (2) and (6):

$$\varphi_n(x)=(2\pi)^{-1/4}(n!)^{-1/2}[(n-p)(n-p-1)]^{-1/2}\exp(-x^2/4)\times[(n-p)He_n(x)$$
$$+(He_p(x)He_{n+1}(x)-He_n(x)He_{p+1}(x))He_{p+1}(x)/J_p(x)],$$
$$n\neq p,(p+1).$$

ıot difficult to study the properties of the second-order Wronskian constructed from ions (16):

$$W_{m,l}(x)\equiv W(u_m,u_l)=f_{m,l}(x)\exp(x^2/2),$$
$$f_{m,l}(x)=q_m(x)q_{l+1}(x)-q_l(x)q_{m+1}(x).$$

ıe derivative of this function we obtain $W'_{m,l}(x)=(l-m)q_m(x)q_l(x)\exp(x^2/2)$. It vs from the definition of the polynomials $q_m(x)$ (17) that their parity coincides with ırity of the polynomial number m, all their coefficients are integer, and $q_m(x)>0$ for m values, and $q_m(x)=0$ only for odd m value and $x=0$, and the zero is simple. It is to establish that $W_{m,l}(0)=l!!(m-1)!!>0$. Hence, it follows that if m and l, both ve, have the different parity then $W_{m,l}(x)$ with $l>m$ decreases for $x<0$, increases >0 and has only one minimum at $x=0$, and $W_{m,l}(x)>0$ is valid for all x values. If then $W_{m,l}(x)$ increases for $x<0$, decreases for $x>0$, and has two symmetrically sed real simple zeros. If m and l have the same parity and we put for definiteness then for odd m,l we have $W'_{m,l}(x)\geq 0$ and the equality holds only for $x=0$ and the is of a second degree, for even m,l we have $W'_{m,l}(x)>0$ and in both cases $x)=W'_{m,l}(-x)$. The function $W_{m,l}(x)$ in this case is an odd monotonically

increasing function having at $x = 0$ the only simple zero for even m and l and triple zero for odd m and l. By this means, for $m = 0, 2, 4, \ldots$ and $l = m + 1, m + 3, m + 5, \ldots$ the functions (16) are suitable for the second order Darboux transformation.

For the new potentials, formula (5) yields the following expression:

$$V_2^{(m,l)}(x) = x^2/4 - 5/2 - 2f''_{m,l}(x)/f_{m,l}(x) + 2[f'_{m,l}(x)/f_{m,l}(x)]^2. \tag{20}$$

Potentials (20) are regular functions on all real axes. The Hamiltonians $H_2^{(m,l)}$ have two additional discrete energy levels, $E = -l - 1$ and $E = -m - 1$, with respect to H_0. The first level corresponds to the ground state and the second one to the excited state. The normalized-at-unity wave functions for these states are

$$\varphi_0^{m,l}(x) = (2\pi)^{-1/4}\sqrt{l!(l-m)}\exp(-x^2/4)q_m(x)/f_{m,l}(x),$$
$$\varphi_1^{m,l}(x) = (2\pi)^{-1/4}\sqrt{m!(l-m)}\exp(-x^2/4)q_l(x)/f_{m,l}(x).$$

Another discrete spectrum wave functions have the form

$$\begin{aligned}\varphi_{n+2}(x) =& (2\pi)^{-1/4}(n!)^{-1/2}[(n+l+1)(n+m+1)]^{-1/2}\exp(-x^2/4)\\ &\times[(n+1)He_n(x) + ((m-l)q_m(x)q_l(x)He_{n+1}(x)\\ &- mlHe_n(x)f_{m-1,l-1}(x))/f_{m,l}(x)], \quad n = 0, 1, 2\ldots.\end{aligned}$$

In full analogy, one can consider higher-order transformations. For example, the fourfold Darboux transformation with two discrete spectrum wave functions with $n = 2$ and $n = 3$ and with two functions (12) with $m = 2$ and $m = 3$ generates the following potential:

$$\begin{aligned}V_4(x) = x^2/4 - 1/2 - 24(1467x^2 + 6x^6 - x^{10})Q^{-1}(x) + 82944\\ \times(105x^2 + 140x^6 + 3x^{10})Q^{-2}(x), \quad Q(x) = 315 + 315x^4 + 9x^8 + x^{12}.\end{aligned}$$

Consider now the generation of simpler potentials of non-elementary form. Choose as a transformation function the following solution of the initial Schrödinger equation:

$$u(x) = e^{x^2/4}(C + \mathrm{erf}(x/\sqrt{2})), \quad H_0u(x) = -u(x).$$

This transformation function generates the well-known family of isospectral potentials [32, 33]

$$V_1(x) = x^2/4 - 3/2 + 2xQ_1^{-1}(x)e^{-x^2/2} + 2Q_1^{-2}(x)e^{-x^2}$$
$$Q_1(x) = \sqrt{\frac{\pi}{2}}(C + \mathrm{erf}(x/\sqrt{2}))$$

(plotted recently in ref. [34]) but the normalizing constant of the ground state function has never been met by the authors in the available literature. For $|C| > 1$ we obtain the following system of normalized-at-unity wave functions

$$\varphi_0(x) = (2\pi)^{-1/4}\sqrt{C^2-1}\cdot u^{-1}(x),$$
$$\varphi_{n+1}(x) = (2\pi)^{-1/4}[(n+1)!]^{-1/2}(He_{n+1}(x)e^{-x^2/4} + He_n(x)Q_1^{-1}(x)e^{-3x^2/4}),$$
$$n = 0, 1, 2, \ldots$$

The general solution of the initial Schrödinger equation for $E=-2$,

$$u_1(x)=\mathrm{e}^{-x^2/4}+x\mathrm{e}^{x^2/4}\left(-C+\sqrt{\frac{\pi}{2}}\operatorname{erf}(x/\sqrt{2})\right),$$

generates another family of the isospectral potentials for $|C|<\sqrt{\pi/2}$:

$$V(x)=-\frac{x^2}{4}-\frac{7}{2}+\frac{1}{2}\left(\frac{x+(2+x^2)(\sqrt{(\pi/2)}\operatorname{erf}(x/\sqrt{2})-C)\mathrm{e}^{x^2/2}}{1-x(C-\sqrt{(\pi/2)}\operatorname{erf}(x/\sqrt{2}))\mathrm{e}^{x^2/2}}\right)^2.$$

Another interesting potential occurs for the single Darboux transformation with the function $\psi_{-1/2}(x)={}_0F_1(3/4,x^4/64)$ (${}_0F_1$ is a conventional symbol for the hypergeometric function ${}_pF_q$), which is a solution of equation (1) for the harmonic oscillator potential with $E=-1/2$. This potential is a double well with $|x_{\min}|\approx 1.68$, $U(x_{\min})\approx -0.94$, and with a maximum at $x=0$ and $U(0)=-1/2$, i.e., the energetic level of the ground state touches the potential curve. It occurs in the discrete spectrum of this potential an additional level $E=-1/2$ with respect to the harmonic oscillator spectrum with the ground state eigenfunction (up to a normalization factor) $\varphi_0(x)=\psi_{-1/2}^{-1}(x)$.

Every potential can be studied in more detail. Consider, for example, potential (17) for $p=2$. The solutions of the new Schrödinger equation are intimately connected with Hermite polynomials. To avoid the square roots in their arguments, we make a change of the variable: $x^2\to 2x^2$. Solutions (18) can be expressed through new nonclassical polynomials $P_n(x)$ orthogonal on all real axes:

$$\varphi_n(x)=N_nP_n(x)R(x)\exp(-x^2/2),R(x)=3+12x^2+4x^4,\quad n=0,1,2,\ldots.$$

The weight function of these polynomials $f(x)=R^{-2}(x)\exp(-x^2)$, in contrast with the classical ones, does not satisfy the Pearson equation. If one normalizes the polynomials $P_n(x)$ with weight $f(x)$ as

$$\int_{-\infty}^{\infty}f(x)P_n^2(x)\mathrm{d}x=1,$$

the normalizing constants are: $N_0^2=24/\sqrt{\pi}$, and $N_{n+1}^2=[2^{n+1}\sqrt{\pi}(n+5)n!]^{-1}$. The connection of the new polynomials with the Hermite ones

$$P_{n+1}(x)=H_n(x)R'(x)+H_{n+1}(x)R(x),\quad P'_{n+1}(x)=2(n+5)R(x)H_n(x),$$

permits one to obtain their generating function

$$F(x,z)=2[4x^5+20x^3+15x-zR(x)]\exp(2zx-z^2)=\sum_{k=0}^{\infty}\frac{P_{k+1}(x)}{k!}z^k,$$

which obeys the differential equation

$$\frac{\partial^2F(x,z)}{\partial z^2}-4(x-z)\frac{\partial F(x,z)}{\partial z}+2[1+2(x-z)^2]F(x,z)=0.$$

In spite of the simple connection of the new polynomials with the Hermite ones, they

satisfy the five-term recursion relations

$$4n(n-1)P_{n-1}(x) - 8xnP_n(x) + 2(1+2n+2x^2)P_{n+1}(x) - 4xP_{n+2}(x) + P_{n+3}(x) = 0.$$

The differential equation for these polynomials follows from those of the Hermite one:

$$P_n''(x) - 2x[1 + 8(3+2x^2)R^{-1}(x)]P_n'(x) + 2(n+5)P_n(x) = 0.$$

The existence of the ladder operators $a = x + \mathrm{d}/\mathrm{d}x$ and $a^+ = x - \mathrm{d}/\mathrm{d}x$ for the harmonic oscillator potential permits one to construct similar operators for the new one. For example,

$$a_L = LaL^+ = x^3 + 7x + 13824x(1+2x^2)R^{-3}(x) - 384x(6-5x^2)R^{-2}(x) - 8x(15-2x^2)R^{-1}(x) + [5 + x^2 + 1152x^2R^{-2}(x) - 24(3-2x^2)R^{-1}(x)] \times \mathrm{d}/\mathrm{d}x - x\mathrm{d}^2/\mathrm{d}x^2 - \mathrm{d}^3/\mathrm{d}x^3.$$

These operators have the properties

$$a_L\varphi_{n+1}(x) = 4n(5+n)\varphi_n(x), \quad a_L^+\varphi_n(x) = 2(n+4)\varphi_{n+1}(x),$$
$$a_L\varphi_0(x) = a_L^+\varphi_0(x) = 0, \quad a_L\varphi_1(x) = 0.$$

It is remarkable that the first polynomial is equal to the unity but the nth one is a polynomial of degree $n+4$. Nevertheless, in agreement with the oscillator theorem for the transformed Hamiltonian the polynomial of order n has n zeros in full real axis. Finally, we cite some of the first polynomials

$$P_0(x) = 1, P_1(x) = 2x(4x^4 + 20x^2 + 15), \quad P_2(x) = 16x^6 + 72x^4 + 36x^2 - 6,$$
$$P_3(x) = 4x(8x^6 + 28x^4 - 14x^2 - 21),$$
$$P_4(x) = 64x^8 + 128x^6 - 480x^4 - 288x^2 + 36,$$
$$P_5(x) = 8x(16x^8 - 216x^4 + 81).$$

3.2 *Effective Coulomb potential*

The next simplest potential, after the harmonic oscillator potential, is the effective Coulomb one:

$$V_0(x) = -2z/x + l(l+1)/x^2, \quad E_n = -z^2/n^2, \quad n = 1,2,3,\ldots,$$
$$\psi_{nl}(x) = N_{nl}^0 x^{l+1}\exp(-zx/n)L_{n-l-1}^{2l+1}(2zx/n),$$
$$N_{nl}^0 = 2/n^2(2/n)^l z^{l+3/2}\sqrt{(n-l-1)!/(n+l)!}.$$

Here, $L_n^\alpha(x)$ are the generalized Laguerre polynomials [30] and functions $\psi_{nl}(x)$ are normalized by the condition $\int_0^\infty \psi_{nl}^2(x)\mathrm{d}x = 1$.

The second order Darboux transformation with these functions with $n = p$ and $n = p+1$ produces the following exactly solvable potentials:

$$V_2^{(l,p,p+1)}(x) = -2z/x + (l+1)(l+4)/x^2 - 2w_0''(x)/w_0(x) + 2(w_0'(x)/w_0(x))^2,$$

where

$$w_0(x) = 2(p+1)L_{p-l-2}^{2l+2}(2zx/p)L_{p-l}^{2l+1}(2zx/(p+1)) + L_{p-l-1}^{2l+1}(2zx/p)[L_{p-l}^{2l+1}(2zx/(p+1)) - 2pL_{p-l-1}^{2l+2}(2zx/(p+1))].$$

The expression for the normalized-at-unity wave functions follows from formulae (6) and (11)

$$\varphi_{nl}(x) = N_{np}\left[\left(\frac{1}{n^2} + \frac{w_1(x)}{np(p+1)w_0(x)}\right)\psi_{nl}(x) - N_{nl}^0 \frac{2(2p+1)}{np(p+1)w_0(x)} \times L_{p-l-1}^{2l+1}(2xz/p)L_{p-l}^{2l+1}(2xz/(p+1))x^{l+1}L_{n-l-2}^{2l+2}(2xz/n)\exp(-zx/n)\right],$$

where

$$w_1(x) = 2n(p+1)L_{p-l-1}^{2l+1}(2xz/p)L_{p-l-1}^{2l+2}(2xz/(p+1)) + L_{p-l}^{2l+1}(2xz/(p+1)) \times [(n-1)L_{p-l-1}^{2l+1}(2xz/p) - 2p(nL_{p-l-2}^{2l+2}(2xz/p)) + L_{p-l-1}^{2l+1}(2xz/p)],$$

and

$$N_{np} = n^2p(p+1)[(p^2-n^2)((p+1)^2-n^2)]^{-1/2}.$$

The simplest nontrivial potentials correspond to the case $l = 0, p = 2$:

$$V_2^{(0,2,3)} = -2z/x + 10/x^2 + 40Q_0^{-1}(xz)(2-xz)/x^2 - 100Q_0^{-2}(xz)(2xz-3)/x^2,$$
$$Q_0(x) = -15 + 10x - 2x^2.$$

The levels with $n = 2$ and $n = 3$ in its discrete spectrum are absent. The normalized-at-unity discrete spectrum eigenfunctions are expressed via a system of new polynomials $P_n(x)$ orthogonal on the interval $(0, \infty)$:

$$\varphi_{nl}(x) = \tfrac{1}{6}N_{nl}^0 N_{n2}\, n^{-2}z^2x^3P_n(xz)Q_0^{-1}(xz)\exp(-xz/n).$$

These polynomials are related with the Laguerre ones as follows:

$$x^3P_n(x) = 10n(-54 + 63x - 22x^2 + 2x^3)L_{n-2}^2(2x/n) + L_{n-1}^1(2x/n)[270n(n-1) - 45(2 - 7n + 5n^2)x + 10(6 - 11n + 5n^2)x^2 - 2(6 - 5n + n^2)x^3].$$

We cite the first five polynomials,

$$P_1(x) = -4, \quad P_2(x) = P_3(x) = 0, \quad P_4(x) = \tfrac{1}{24}(-84 + 63x - 18x^2 + 2x^3)$$
$$P_5(x) = \tfrac{8}{625}(-875 + 700x - 220x^2 + 30x^3 - x^4).$$

The properties of $P_n(x)$ strongly differ from the properties of the Laguerre polynomials, though they are completely defined by the latter. For example, for $n > 3$ the polynomial $P_n(x)$ is a polynomial of degree n but it has $n-3$ zeros in the $(0, \infty)$ interval, and this fact corresponds well to the oscillator theorem for the Hamiltonian H_1.

$l = 0$, $n_1 = 3$, $n_2 = 4$, $n_3 = 5$ and $n_4 = 6$, the new potential is

$$V_4^{(0,3,4,5,6)}(x) = -2z/x + 38/x^2 - 144x^{-2}Q_2^{-1}(xz)(492 - 228xz + 27x^2z^2 - x^3z^3) - 31104x^{-2}Q_2^{-2}(xz)(9240 - 2640xz + 255x^2z^2 - 8x^3z^3),$$
$$Q_2(x) = 11880 - 3960x + 540x^2 - 36x^3 + x^4.$$

The levels with $n = n_1, n_2, n_3$, and n_4 are absent in its discrete spectrum.

In contrast with the harmonic oscillator potential, we can obtain here isospectral family of potentials of elementary form. For this purpose, the solutions of the initial Schrödinger equation singular at $x = 0, \infty$ can be used as transformation functions. We can construct elementary solutions having such properties with the use of the following functions:

$$u_{nl}(x) = x^{-l}L_n^{-2l-1}(2zx/(n-l))\exp(-zx/(n-l)),$$
$$H_0u_{nl}(x) = -z^2/(n-l)^2u_{nl}(x).$$

For $l = 1$ and $n = 2$, the general solution of the Schrodinger equation has the form

$$u(x) = x^{-1}e^{-zx}Q_3(xz), Q_3(x) = 1 + 2x + 2x^2 + Ce^{2x}, \quad H_0u(x) = -z^2u(x).$$

This transformation function leads to the family of isospectral potentials with a hydrogen-like discrete spectrum:

$$V_1(x) = -2z/x - 16z^3x(xz-1)Q_3^{-1}(xz) + 32z^6x^4Q_3^{-2}(xz),$$
$$C \in (-\infty, -1) \cup (0, \infty).$$

These potentials with $z = 1$ are plotted in figure 2. The normalized-at-unity eigenfunctions are

$$\varphi_1(x) = 2\sqrt{z^3C(C+1)}u^{-1}(x),$$
$$\varphi_n(x) = N_{n1}^0nz^{-1}(n^2-1)^{-1/2}[x(2-xz/n)L_{n-2}^3(2xz/n) - 2x^2z/nL_{n-3}^4(2xz/n) + x(1+xz+Ce^{2xz}(1-xz)+2x^3z^3)L_{n-2}^3(2xz/n)Q_3^{-1}(xz)]\exp(-xz/n),$$
$$n = 2, 3, \ldots.$$

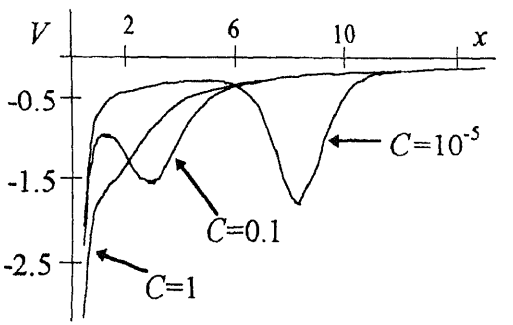


Figure 2. The isospectral Coulomb-like potentials.

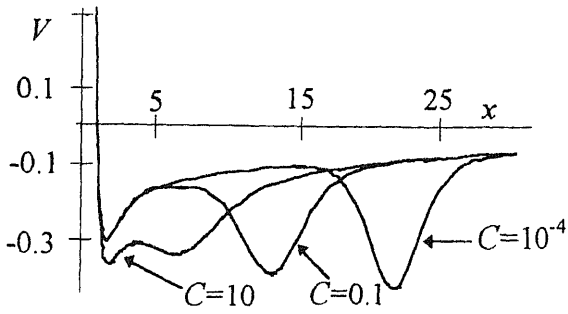


Figure 3. The isospectral Coulomb-like potentials with a centripetal member.

For $n = 4,\ l = 2$ the function

$$u(x) = x^{-2}\exp(-xz/2)Q_4(xz), \quad Q_4(x) = 24 + 24x + 12x^2 + 4x^3 + x^4 + Ce^x$$

produces a family of isospectral potentials with a centripetal member

$$V_1(x) = -2z/x + 2/x^2 + 2z^5x^3(4 - xz)Q_4^{-1}(xz) + 2z^{10}x^8Q_4^{-2}(xz)$$

plotted for $z = 1$ in figure 3. The normalized-at-unity wave functions for these potentials are given as follows:

$$\varphi_2(x) = z^{5/2}\sqrt{C(C/4! + 1)}\cdot u^{-1}(x), C \in (-\infty, -4!) \cup (0, \infty)$$

$$\varphi_n(x) = 2N_{n2}^0 n/z(n^2 - 4)^{-1/2}\exp(-xz/n)[x^2(3 - xz/n)L_{n-3}^5(2xz/n) - 2zx^3/n$$
$$\times L_{n-4}^6(2xz/n) + \tfrac{1}{2}x^2Q_4^{-1}(xz)(96 + 72xz + 24x^2z^2 + 4x^3z^3 + x^5z^5$$
$$+ (4 - xz)Ce^{xz})L_{n-3}^5(2xz/n)], n = 3, 4, 5, \ldots.$$

It is interesting to note that, as can be seen from figures 2 and 3, the plots of the Coulomb-like potentials resemble the diagram of a moving soliton, especially if one changes the sign: $V \to -V$.

Note as well that the Jost functions for every potential can be found with the help of the asymptotic behavior of the transformation functions. We do not dwell on these calculations.

3.3 *Morse potential*

One more example is a rather complicated potential known as the Morse potential:

$$V_0(x) = \exp(-2\alpha x) - A\exp(-\alpha x),$$

$$E_n = -\tfrac{1}{4}[A - \alpha(2n + 1)]^2, \psi_n(x) = z^\mu\exp(-z/2)\cdot L_n^{2\mu}(z),$$

$$z = \frac{2}{\alpha}\exp(-\alpha x), \mu = \sqrt{|E_n|}/\alpha, \quad n = 0, 1, 2, \ldots.$$

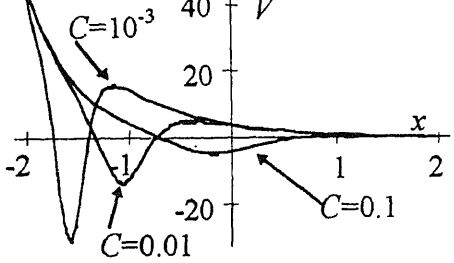


Figure 4. The isospectral one-level Morse potentials with $k = 2$ and $\alpha = 1$.

We consider the variable x to span all real axis. The second order Darboux transformation with the function $\psi_1(x)$ and $\psi_2(x)$ gives a new potential of the form

$$V_2^{(1,2)}(x) = \exp(-2\alpha x) - \exp(-\alpha x)(A - 4\alpha) + 128(2\alpha - A)^{-1}\alpha^3 Q_4^{-2}(x)$$
$$\times [(A - 3\alpha)\exp(\alpha x) - 1] + 8(2\alpha - A)^{-1}\alpha^2 Q_4^{-1}(x)$$
$$\times [4\alpha - (A^2 - 5A\alpha + 6\alpha^2)\exp(\alpha x),$$
$$Q_4(x) = 4 - 4(A - 3\alpha)\exp(\alpha x) + (A^2 - 5A\alpha + 6\alpha^2)\exp(2\alpha x), \quad A \neq 2\alpha.$$

In its discrete spectrum, the states with $n = 1$ and $n = 2$ are missing.

On the base of this potential, we can construct a family of isospectral potentials of elementary form with a single discrete spectrum energy level. This possibility is due to the fact that when μ is entire or half-integer, the general solution of the Schrödinger equation has an elementary form. Let μ take a fixed value as $\mu = -(p+1)/2$ and $A = -p\alpha, p = 0, 1, 2, \ldots$. In this case, the Morse potential

$$V_0(x) = \tfrac{1}{4}\alpha^2 z^2 + \tfrac{1}{2}\alpha^2 pz$$

has no discrete spectrum at all. However, for $E_0 = -\frac{1}{4}(p+1)^2\alpha^2$ we have the following general solution of the initial Schrödinger equation:

$$u(x) = e^{-z/2} z^{-(p+1)/2} Q_5(z), \quad Q_5(z) = 1 + Ce^z \left(z^p + \sum_{i=1}^{p} (-1)^i p!/(p-i)! z^{p-i} \right).$$

This function when used as a transformation function, generates a one-level potential

$$V_1(x) = \tfrac{1}{4}\alpha^2 z^2 + \alpha^2(p/2 + 1)z - 2\alpha^2 z \mathrm{d}\log Q_5(z)/\mathrm{d}z - 2\alpha^2 z^2 \mathrm{d}^2 \log Q_5(z)/\mathrm{d}z^2,$$

with the energy level $E = E_0$ and the normalized-at-unity wave function of the form:

$$\psi_0(x) = [\alpha C((-1)^p p! C + 1)]^{1/2} u^{-1}(x).$$

It is clear from this relation that for even p values for the constant C we have $C \in (-\infty, -1/p!) \cup (0, \infty)$ and for odd p values it must fall in the interval $C \in (0, 1/p!)$.

The simplest of these potentials corresponds to $p = 0$:

$$V_1(x) = \exp(-2\alpha x) + 2\alpha\left[1 - C\exp\left(\frac{2}{\alpha}\,\mathrm{e}^{-\alpha x}\right)\right]$$

$$\times\, Q_6^{-1}(x) - 8C\exp\left(\frac{2}{\alpha}\mathrm{e}^{-\alpha x}\right)Q_6^{-2}(x),$$

$$Q_6(x) = \mathrm{e}^{\alpha x}\left(1 + C\exp\left(\frac{2}{\alpha}\mathrm{e}^{-\alpha x}\right)\right).$$

These potentials are plotted for $p = 2$ in figure 4.

Acknowledgement

This research paper has been partially supported by RFBR grant No. 97–02–16279.

References

[1] G Darboux, *Lecons sur la theorie generale des surfaces et les application geometriques du calcul infinitesimal. Deuxiem partie* (Gauthier Villars et fils, Paris, 1889)
[2] E L Ince, *Ordinary differential equations* (Dover, New York, 1926)
[3] E Schrödinger, *Proc. R. Irish Acad.* **A47**, 53 (1941)
[4] L Infield and T E Hull, *Rev. Mod. Phys.* **23**, 19 (1951)
[5] E Witten, *Nucl. Phys.* **A188**, 513 (1981)
[6] F Cooper, A Khare and U Sukhatme, *Phys. Rep.* **251**, 267 (1995)
[7] V Matveev and A Salle, *Darboux transformations and solitons* (Springer, Berlin, 1991)
[8] A P Veselov and A B Shabat, *Functionalnii analiz i ego prilozhenia.* **27**, 1 (1993)
[9] A Degasperis and A Shabat, *Teor. Mat. Fiz.* **100**, 230 (1994)
[10] B G Idlis, M M Musahanov, M Sh Usmanov, *Teor. Mat. Fiz.* **101**, 47 (1994)
[11] V G Bagrov and B F Samsonov, *Teor. Mat. Fiz.* **104**, 356 (1995)
[12] M Crum, *Quart. J. Math.* **6**, 263 (1955)
[13] M G Krein, *Dokl. Akad. Nauk SSSR* **113**, 970 (1957)
[14] A A Andrianov, M V Ioffe and V P Spiridonov, *Phys. Lett.* **A174**, 273 (1993)
[15] F Calogero and A Degasperis, *Spectral Transform and Solitons: Tools to solve and investigate nonlinear evolution equations* (North Holland Publ. Co., Amsterdam, 1982) vol. 1
[16] B F Samsonov, *J. Phys.* **A28**, 6989 (1995)
[17] V G Bagrov, I N Ovcharov and B F Samsonov, *J. Moscow Phys. Soc.* **5**, 191 (1995)
[18] V Adler, *Teor. Mat. Fiz.* **101**, 323 (1994)
[19] B F Samsonov, *Mod. Phys. Lett.* **A11**, 1563 (1996)
[20] K Shadan and P C Sabatier, *Inverse problems in quantum scattering theory* (Springer, Berlin, 1977)
[21] C V Sukumar, *J. Phys.* **A18**, 2937 (1985)
[22] M M Nieto, *Phys. Lett.* **B145**, 208 (1984)
[23] C V Sukumar, *J. Phys.* **A18**, 2917 (1985)
[24] V P Berezovoi and A I Pashnev, *Teor. Math. Fiz.* **70**, 148 (1987)
[25] A Khare and U Sukhatme, *J. Phys.* **A22**, 2847 (1989)
[26] W-Y Keung, U Sukhatme, Q Wang and T Imbo, *J. Phys.* **A22**, 987 (1989)
[27] Cao xuan Chuan, *J. Phys.* **A24**, L1155 (1991)
[28] L D Faddeev, *Usp. Mat. Nauk.* **14**, 57 (1959)
[29] G A Natanson, *Teor. Mat. Fiz.* **38**, 219 (1979)
[30] *Handbook of Mathematical Functions* edited by M Abramowits and I A Stegun (National Bureau of Standards) *Appl. Math.* Series 55 (1964)

[31] Iu S Dubov, V M Eleonskii and N E Kulagin, *Zh. Eksper. Theor. Fiz.* **102**, 814 (1992)
[32] P B Abraham and H E Moses, *Phys. Rev.* **A22**, 1333 (1980)
[33] B Mielnik, *J. Math. Phys.* **25**, 3387 (1984)
[34] D J Fernandez, V Hussin and L M Nieto, *J. Phys.* **A27**, 3547 (1994)

Detection of determinism and randomness in time series: A method based on phase space overlap of attractors

M S GOPINATHAN
Department of Chemistry, Indian Institute of Technology, Chennai 600 036, India
Email: msgopi@hotmail.com

MS received 19 March 1997; revised 25 August 1997

Abstract. The space overlap of an attractor reconstructed from a time series with a similarly reconstructed attractor from a random series is shown to be a sensitive measure of determinism. Results for the time series for Henon, Lorenz and Rössler systems as well as a linear stochastic signal and an experimental ECG signal are reported. The overlap increases with increasing levels of added noise, as shown in the case of Henon attractor. Further, the overlap is shown to decrease as noise is reduced in the case of the ECG signal when subjected to singular value decomposition. The scaling behaviour of the overlap with bin size affords a reliable estimate of the fractal dimension of the attractor even with limited data.

Keywords. Chaos; determinism; randomness; phase space overlap.

PACS No. 05.45

A basic problem in the analysis of a nonlinear dynamical system based on its measured time series is how to decide whether the dynamics is generated by a deterministic or a random process or by a combination of these two. The first step in the study of an experimental time series is usually the reconstruction of the attractor by the standard procedure using time delay vectors [1, 2] . Methods available to disentangle deterministic chaos from a stochastic process can be broadly divided into two categories: (a) Study of the temporal properties of the flows on the attractor. These include determination of Lyapunov exponents [3] measuring the divergence of trajectories, the average tangent of the flows on the attractor [4–6] and short term predictability of the time series [7]. (b) Methods based on the spatial structure of the attractor such as correlation dimension [8] and K_2 entropy [9]. Presence of even low levels of noise, which is unavoidable in a measured time series, can seriously affect the efficiency of these methods. For instance, determination of Lyapunov exponents and correlation dimension is unreliable [10–12] when noise level exceeds 1%.

Here we propose a simple geometric procedure to determine the overlap, η, in phase space of the reconstructed chaotic attractor under study and a surrogate random attractor reconstructed in the same manner, with identical reconstruction parameters, from a series of random numbers with a chosen distribution. It turns out that the magnitude of the overlap varies from $\eta \approx 1$ for a purely random process to $\eta \approx 0.1$ for deterministic chaos. η is found to be very sensitive especially to low levels of noise added to a chaotic time

increases with additive noise and the attractor for a pure random series fills up the whole phase space.

In what follows, we first detail the procedure for the evaluation of the overlap η and then apply it to time series generated for standard chaotic systems like the Lorenz, Rössler and Henon systems [13] and to a linear stochastic signal [14]. An experimental ECG signal for a healthy human heart is also analysed. We then report the effect on η due to addition of different levels of noise to the Henon attractor as a test case. The scaling behaviour of η is also investigated and it is shown to give a fair estimate of the fractal dimension d_f of the attractor even with a limited number of data points. The sensitivity of η and d_f to noise is demonstrated.

The following procedure is used to determine the phase space overlap η of the attractor reconstructed from the time series with an attractor reconstructed similarly from a random series.

The time series $Y = (y_1, y_2, \ldots, y_M)$ of M data points is first linearly scaled so as to lie in a unit interval as

$$x_i = [(y_i - y_{\min})/(y_{\max} - y_{\min})] - 1/2. \tag{1}$$

Thus we have $-1/2 \leq x_i \leq 1/2$. This scaling fixes the volume of the state space in which the attractor is embedded to be unity, for convenience of comparison with other attractors. The attractor for this time series is then reconstructed by the method of delays [2] using delay vectors, $\mathrm{s}_i = \{x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(d-1)\tau}\}$, choosing appropriate values of reconstruction parameters namely delay time τ and embedding dimension d. This gives $N = M - (d-1)\tau$ points on the attractor. We call this attractor $\mathbf{S} = \{\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_N\}$.

The same reconstruction procedure with identical τ and d values is now applied to M computer generated random points of uniform distribution and a random attractor of N points is thus constructed within the same unit volume of phase space defined by the attractor $\mathbf{S}$. We designate the random attractor as $\mathbf{R} = \{\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_N\}$.

As an example, figure 1 shows the superposition of the reconstructed Rössler attractor and the corresponding random attractor for $N = 5000$ points, $\tau = 5$ and $d = 3$ as a typical case. The near coverage of the unit phase space by the random attractor and the limited region to which the Rössler attractor is confined may be noted.

We now define the overlap η of the two attractors $\mathbf{S}$ and $\mathbf{R}$ as

$$\eta(N, \varepsilon) = \frac{1}{N} \sum_{r_i \in R}^{N} [1 - f(g_{r_i})], \tag{2a}$$

where $f(z) = 0$ for $z \neq 0$ and $f(0) = 1$, and

$$g_{r_i} = \sum_{s_j \in S}^{N} \Theta(\varepsilon - |\mathrm{r}_i - \mathrm{s}_j|). \tag{2b}$$

Here Θ is the Heaviside step function. ε is the chosen bin size. The summation in (2a) goes over the N points of the random attractor $\mathbf{R}$ and that in (2b) goes over the N points of the attractor $\mathbf{S}$ under study. That η as defined is a measure of the overlap between two attractors in their common phase space can be seen as follows.

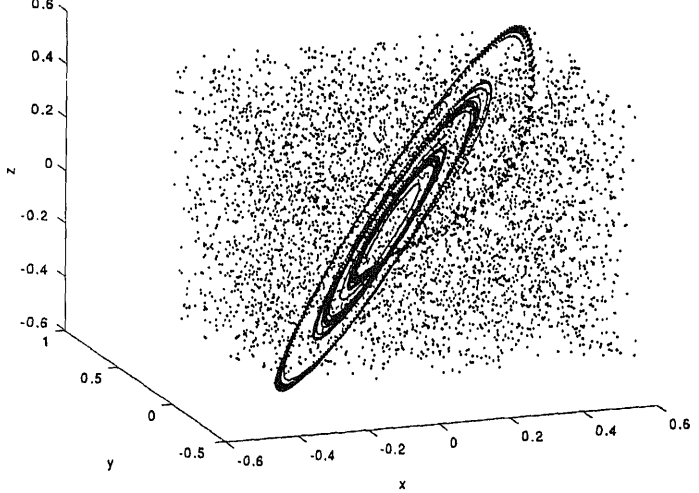


Figure 1. Rössler attractor reconstructed from x time series of the system superimposed on a random attractor of the same number of data points, $N = 5000$. x series is generated by 4th order Runge–Kutta integration of the Rössler system: $\{\dot{x}, \dot{y}, \dot{z} = -(y+z), x + ay, b + xz - cz\}$ with $a = 0.2$, $b = 0.2$ and $c = 5.7$, and the uniformly distributed random series is generated by the Matlab language command '*rand*'. Reconstruction parameters for both attractors are: dimension $d = 3$ and time delay $\tau = 5$.

Consider the two atrractors R and S to be identical. Thus for every point r_i, there is an identical point s_i so that $g_{r_i} > 0$ and hence $f(g_{r_i}) = 0$ for every r_i. Consequently, the sum in (2b) yields N and we therefore have $\eta = 1$ for this case of two identical attractors.

On the other hand, if the two attractors occupy non-overlapping regions of the unit phase space volume such that no point of R is within a distance ε of any point of S, we have $g_{r_i} = 0$ for every r_i. Therefore, $f(g_{r_i}) = 1$ for every r_i and we have $\eta = 0$ for this case. It is in this sense that we call η as the overlap between two attractors and its value can range from 1 to 0. η, as per eq. (2a) is the ratio of the number of points on the random attractor R that are within a distance ε of the points on the chaotic attractor S, to the total number of points N. Since the random attractor R is supposed to fill the whole of the unit phase space, η can be considered to be a measure of the spatial extent of the chaotic attractor S. This procedure is essentially a Monte Carlo [15] method of determining the volume of an object in a multidimensional space.

The overlap η should be distinguished from the correlation integral $C(\varepsilon)$ defined as [8]

$$C(\varepsilon) = \frac{1}{N^2} \sum_{i,j} \Theta(\varepsilon - |\mathbf{x}_i - \mathbf{x}_j|) \tag{3}$$

or the cross correlation sum [16]

$$C_{XY}(\varepsilon) = \frac{1}{|X|\,|Y|} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in Y} \Theta(\varepsilon - |\mathbf{x} - \mathbf{y}|). \tag{4}$$

between two attractors that fall within a distance of ε.

Ideally $\eta(N, \varepsilon)$ should be studied as a function of N and ε. However we may make use of the assumption that the N random points uniformly cover the whole of the unit phase space volume so that we may fix ε for a given N as

$$\varepsilon^d = 1/N, \tag{5}$$

where d is the integer phase space dimension. From now on we denote η as $\eta(N)$ or as $\eta(\varepsilon)$ on account of the relation (5). This relation is prompted by the following reasoning. ε^d is the volume of a single bin and since we have scaled the phase space to unit volume, this should be given by $1/N_b$ where N_b is the number of bins. If N_b were chosen to be $\gg N$, corresponding to very small value of ε then $\eta \approx 0$ since the bin is so small that the probability of a given bin having two points, one each from the two attractors, would be very small. On the other hand, if $N_b \ll N$, corresponding to large ε, then $\eta \approx 1$ since every bin is likely to have a point from each of the two attractors. Hence setting $N_b = N$ is a reasonable choice, fixing the bin volume to the effective volume occupied by a point

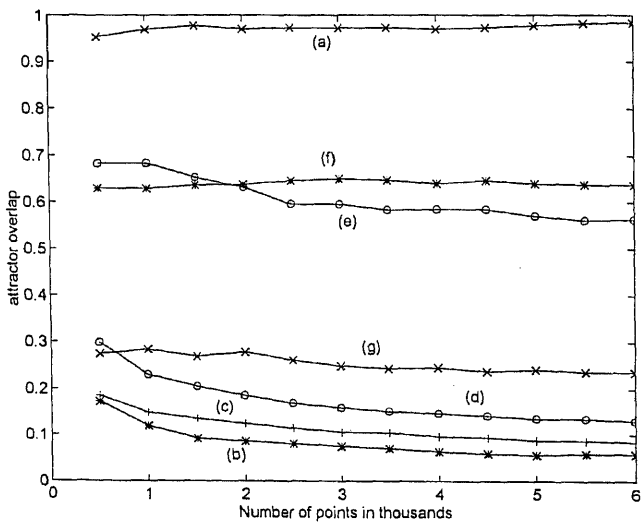


Figure 2. Overlap η (eq. 2) of reconstructed attractor with the reconstructed random attractor, as a function of data points used in reconstruction. For each curve, the same reconstruction parameters, namely phase space dimension d and time delay τ are used for the two attractors. (**a**) Another random attractor, $d = 3, \tau = 10$; (**b**) Rössler x series with parameters specified as in figure 1, $d = 3, \tau = 5$; (**c**) Lorenz x series of the system $\{\dot{x}, \dot{y}, \dot{z} = -\sigma x + \sigma y, rx - y - xz, -bz + xy\}$ with $\sigma = 10$, $b = 8/3$, $r = 28$, $d = 3$, $\tau = 5$; (**d**) Clean Henon map: $\{x_{k+1}, y_{k+1} = y_k + 1 - ax_k^2, bx_k\}$ with $a = 1.4$ and $b = 0.3, d = 2, \tau = 1$; (**e**) Stochastic time series of eq. (6), $d = 2, \tau = 5$; (**f**) x series of noisy Henon map, eq. (7) , with $p = 50\%$, $d = 2, \tau = 5$; (**g**) Experimental ECG signal from a healthy human heart recorded at sampling frequency of 250 Hz with 12 bit resolution, $d = 2, \tau = 4$. In order to avoid minor fluctuations in η due to the random number generation algorithm, the random numbers are chosen from a very large set of random numbers ($\sim$ 15000) generated just once. The time delay τ used for the various attractors are rather arbitrary and no optimization of τ has been done in this work.

on the random attractor. This procedure is similar to the usual fixing of ε in the intermediate region where the logarithm of the correlation integral of (3) scales linearly with $\log(\varepsilon)$ [13].

$\eta(N)$ is plotted against N in figure 2 for a few dynamical systems, for N going from 500 to 6000 points. Curve (a) is for two independently realised random attractors in $3d$. It shows that $\eta(N) \approx 1$ for practically all values of N. Curves (b) and (c) are for the Rössler and Lorenz differential systems reconstructed in $3d$ and curve (d) is for the reconstructed discrete Henon attractor in $2d$. Remarkably, for all these attractors representing deterministic chaos, $\eta \approx 0.1$, for large N. Further, $\eta(N)$ levels off as $N \rightarrow 6000$ in all these cases. The dependence of η on N or equivalently on ε and its relation to the fractal dimension of the attractor is discussed later.

A time series generated by a linear stochastic process has been reported earlier [14] to correspond to a correlation dimension ≈ 2.5, a value normally considered indicative of low dimensional chaos. This series thus constitutes a counter-example to the trend of correlation dimension increasing with increasing embedding dimension for stochastic processes. The series is

$$\dot{x} = \theta x(t) + \omega(t), \tag{6}$$

where $\theta = -0.9$ and $\omega(t)$ is a standard Gaussian white noise process. Our estimate of η for the attractor for this series reconstructed in $2d$ is 0.56, clearly showing its stochastic origin (see curve (e) in figure 2).

Reported in figure (2), curve (f), is another noisy series generated by adding uniform random noise to the Henon map as

$$x'_n = x_n + (p/100)R_p l_x; \quad y'_n = y_n + (p/100)R_q l_y, \tag{7}$$

where (x_n, y_n) is the nth iterate of the clean Henon map and the primed coordinates denote the same with p per cent noise added. R_i is a random number with uniform distribution lying in the unit interval between -0.5 and $+0.5$ and l_x and l_y are the maximum extent of the clean Henon map in the x direction: $l_x = |x_{\max} - x_{\min}|$ and similarly for l_y. Figure 3, which is a superposition of the clean Henon attractor with one of 10% noise, clearly shows the spreading of the attractor with noise. Curve (f) in figure 2 is for the reconstructed Henon attractor with 50% noise and this gives an η value of 0.63 for large N, representing the high level of noise.

As a final example, the $\eta(N)$ values for an experimental time series, namely the ECG of a healthy human heart is displayed by curve (g) in figure 2. The large N value of η for this series is 0.23 which indicates deterministic chaos probably mixed with some noise. The presence of noise can be verified by applying one of the noise reduction techniques [10] to the time series and then determining η. Presently we applied the Broomhead and King's method [17] of singular value analysis to the ECG signal and reconstructed the attractor using the time series obtained by a projection of the trajectory matrix onto the most significant singular vector. This reconstructed attractor gives a considerably reduced η value of 0.16 indicating the presence of noise in the original data which corresponded to $\eta = 0.23$.

A study of η for the Henon map of (7), for a fixed value of $N = 5000$, at various values of the percentage of added noise, p, is reported in figure 4. The high sensitivity of η to noise, at low levels of noise is obvious. Note that η is proposed here as an index of

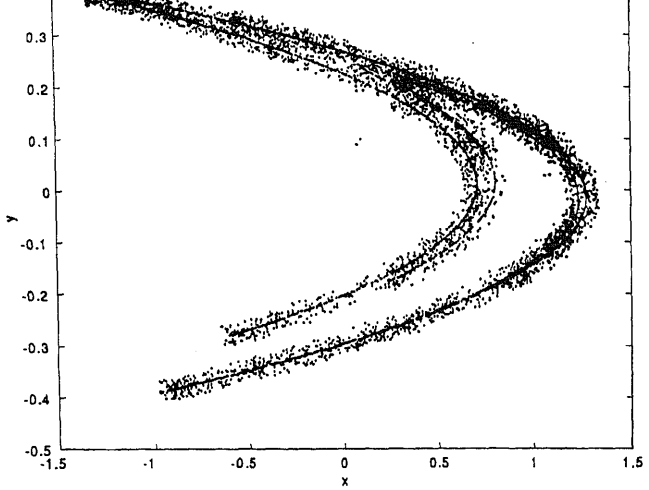


Figure 3. Clean Henon attractor as defined in figure 2, superimposed on another reconstructed noisy Henon attractor with $p = 50\%$ i.e. noise/signal ratio of 0.5. Number of data points = 5000, $d = 2$ and $\tau = 1$ for both attractors.

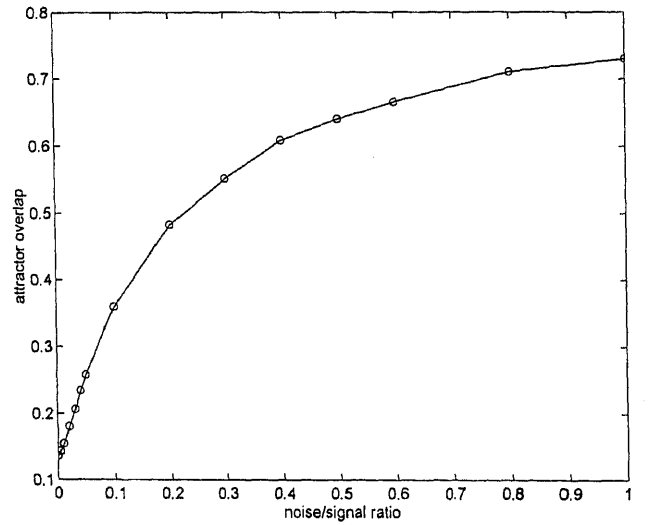


Figure 4. Overlap η of noisy Henon attractor with random attractor as a function of added noise. Noise to signal ratio corresponds to $p/100$ (eq. 7).

determinism in the time series and hence if noise is present in the observed signal, it should show up as increased value of η compared to that of a noise free signal. A major problem in the analysis of experimental signals is to discern if the observed nonlinearity is part of the dynamics or due to noise. Hence removal or addition of noise should alter the index of determinism (η in our case) sensitively, especially at low levels of noise

which is the level normally present in a measured signal. At high levels of noise, η is anyway high, approaching 1, indicating the signal is dominantly noise. Figure 4 shows that η has this desirable property.

These numerical results indicate that the overlap η of a chaotic attractor with its surrogate random attractor in phase space can serve as one of the measures of stochasticity in a time series. η values of the order of ≈ 0.1 is indicative of deterministic chaos and significantly larger values point towards the presence of noise.

It is possible to determine, at least approximately, the fractal (box counting) dimension of the attractor from $\eta(\varepsilon)$ values determined as a function of ε. Denoting the number of points on the random attractor that fall on the chaotic attractor as N_a, we may set

$$N_a = v/\varepsilon^{d_f}, \tag{8}$$

where v is the constant volume occupied by the attractor in the total phase space of unit volume, after the transients have died down and d_f is the fractal dimension of the attractor. Combining (5) and (8) we have

$$\eta(\varepsilon) = \frac{N_a}{N} \propto \varepsilon^{d-d_f}. \tag{9}$$

Plots of $\log\eta$ vs. $\log\varepsilon$ for several attractors are given in figure 5. From the linear least squares fit (solid lines), it is clear that a scaling region exists for all these attractors, despite the different dimensions and time delay used, for moderate number of points in the range 500–6000. From the slope of these lines, $(d - d_f)$ and hence the fractal dimension d_f can be determined. This gives values of 2.04 ± 0.02 and 1.34 ± 0.01 for the Lorenz and Henon attractors for the chosen parameters, which compares well with the

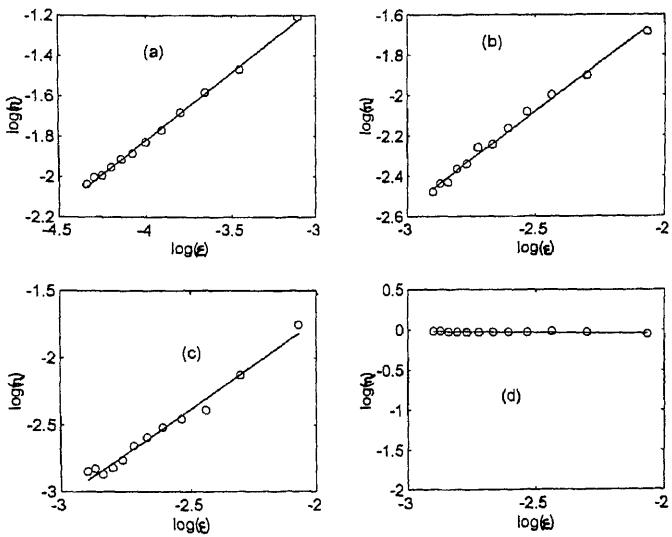


Figure 5. Log(η) vs. log(ε) for various attractors, for $N = 500$ to 6000 in steps of 500. (**a**) Henon attractor; (**b**) Lorenz attractor; (**c**) Rössler attractor; (**d**) random attractor. All time series and attractor reconstruction parameters are as defined in figure 2.

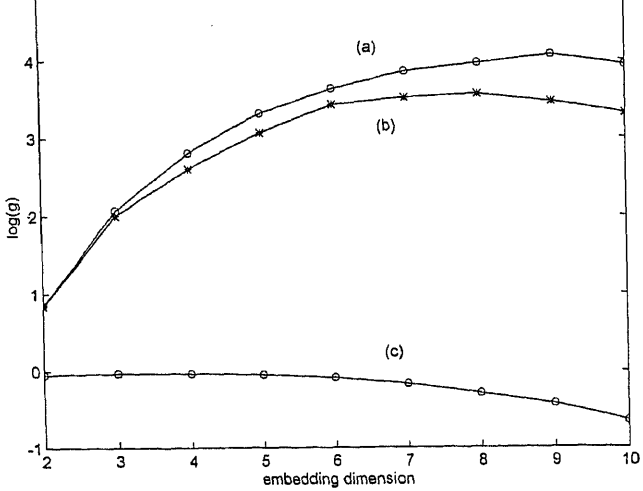


Figure 6. $\mathrm{Log}(\eta/\varepsilon^{d-d_f})$ vs. embedding dimension d, (**a**) for Henon attractor with $N = 5000$; (**b**) Henon attractor with $N = 1000$; (**c**) random attractor with $N = 2000$. $d_f = 1.34$ for the Henon attractor and $d_f = d$ for the random attractor. Attractor and reconstruction parameters as in figure 2.

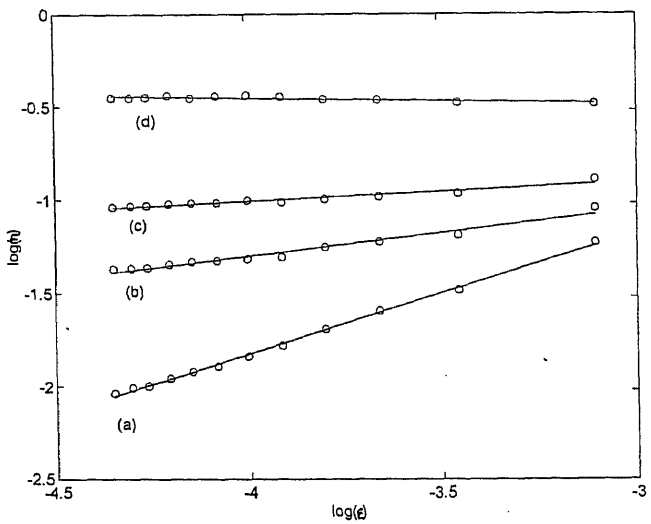


Figure 7. $\mathrm{Log}(\eta)$ vs. $\log(\varepsilon)$ for Henon attractor with different noise levels, p, eq. (7). (**a**) Clean Henon, $p = 0\%$; (**b**) $p = 5\%$; (**c**) $p = 10\%$; (**d**) $p = 50\%$. Attractor and reconstruction parameters as in figure 2.

exact values 2.05 ± 0.01 and 1.21 ± 0.01 respectively [18]. For the Rössler attractor we get $d_f = 1.67 \pm 0.06$ for the parameters chosen here. For the random attractor, as expected, $(d - d_f) \approx 0$ showing that η is nearly independent of ε in this case (see figure 5(d)).

From (9), the ratio $g = \eta(\varepsilon)/\varepsilon^{d-d_f}$ should be a constant. A plot of $\log(g)$ as a function of the embedding dimension is shown in figure 6 for the Henon map, for fixed number of data points $N = 5000$ and $N = 1000$. Saturation of this ratio around $d \approx 6-7$ irrespective of N is clear. On the other hand, for a random attractor, as $d - d_f \approx 0$, g is nearly independent of the embedding dimension as seen from curve (c) in figure 7.

In order to study the effect of added noise on the attractor overlap, plots of $\log \eta$ vs. $\log \varepsilon$ for the Henon attractor with various levels of noise along with the best linear fit are displayed in figure 7. The systematic decrease in slope with increase in noise level is obvious. The slopes are 0.66, 0.26, 0.12 and -0.02 for levels of noise, $p = 0, 5, 10$ and 50% respectively corresponding to increasing fractal dimensions of 1.34, 1.74, 1.88 and 2.02 respectively.

It is also interesting to note that the quality of the linear fit as measured by the correlation coefficient deteriorates with increasing noise, these being 0.99, 0.98, 0.97 and -0.60 for the above p values. This demonstrates that the estimation of the dimension in the presence of very high noise levels of the order of 50% or more is unreliable.

In summary, we have proposed that the phase space overlap of reconstructed attractor with its random surrogate attractor can serve as an index of determinism. The low value of the order of ≈ 0.1 for the phase space overlap η of a deterministic chaotic attractor with its random attractor is found to be true for a variety attractors studied here, irrespective of the different phase space dimensions and time delay employed in the reconstruction of the attractors from time series. η is found to be extremely sensitive to noise, especially at low levels of noise. Reduction in η can also serve as a measure of the efficiency of noise-reduction technique applied to the time series. The existence of a scaling region for moderate number of points of a few thousands makes possible the economical determination of the fractal dimension of the attractor.

Further studies of the overlap method using surrogate data generated, for example by randomizing the phase of the Fourier transform of the time series [19], will be reported in future.

References

[1] F Takens, in *Dynamical Systems and Turbulence* of 'Springer Lecture Notes in Mathematics' edited by D A Rand and L-S Young (Springer-Verlag, New York, 1981) vol. 898, p. 366
[2] J-P Eckmann and D Ruelle, *Rev. Mod. Phys.* **57**, 617 (1985)
[3] A Wolf, J B Swift, H L Swinney and J A Vastano, *Physica* **D16**, 285 (1985)
[4] D T Kaplan and L Glass, *Phys. Rev. Lett.* **68**, 427 (1972)
[5] D T Kaplan and L Glass, *Physica* **D64**, 431 (1993)
[6] L W Salvino and R Cawley, *Phys. Rev. Lett.* **73**, 1091 (1994)
[7] D H Holton and R M May, in *The nature of Chaos* edited by T Mullin (Clarendon Press, Oxford, 1993) p. 149
[8] P Grassberger and I Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983)
[9] P Grassberger and I Procaccia, *Phys. Rev.* **A28**, 2591 (1983)
[10] E J Kostelich and T Schreiber, *Phys. Rev.* **E48**, 1752 (1993)
[11] T Schreiber, *Phys. Rev.* **E48**, R13 (1993)
[12] However, a recent report shows that for IID Gaussian noise, correlation dimension can be reliably estimated up to noise levels of 20%. See C Diks *Phys. Rev.* **E53**, R4263 (1996)
[13] S H Strogatz, *Nonlinear dynamics and chaos* (Addison-Wesley Publishing Co., Reading, 1994)

[14] A Provenzale, L A Smith, R Vio and G Murante, *Physica* **D58**, 31 (1992)
[15] W H Press, S A Teukolsky, W T Vetterling and B P Flannery, *Numerical recipes in Fortran* (Cambridge University Press, 1992) p. 295
[16] H Kantz, *Phys. Rev.* **E49**, 5091 (1994)
[17] D Broomhead and G P King, *Physica* **D20**, 217 (1986)
[18] P Grassberger and I Procaccia, *Physica* **D9**, 189 (1983)
[19] J Theiler, A Longlin, B Galdrikin and J D Farmer, *Physica* **D58**, 77 (1992)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 591–601

Schrödinger picture formalism of Φ^6 theory

K P SATHEESH and K BABU JOSEPH
Department of Physics, Cochin University of Science and Technology, Kochi 682 002, India

MS received 24 July 1997

Abstract. The static effective potential for a scalar field with Φ^6 interaction is calculated using the effective action in Schrödinger picture formalism. It is found that the effective potential obtained is same as the Gaussian effective potential as far as static case is concerned. Equivalence with the CJT formalism can also be established. As in CJT formalism after renormalization an unrenormalized mass term persists. Nonzero turning points are obtained both for positive and negative λ. Results are analysed numerically. Graphical analysis indicates a behaviour similar to that obtained for CJT formalism at zero temperature.

Keywords. Gaussian trial function; variational principle; Schrödinger picture; static effective potential; effective action; renormalization; CJT formalism; Gaussian effective potential.

PACS Nos 11.10; 11.15

1. Introduction

Variational methods in functional Schrödinger picture have been found to be very useful in the study of detailed structures of the quantum fields, both for the bosonic and fermionic field theories. Recently various interesting applications of this formalism have been presented [1–3]. In $(2+1)$ dimensional Thirring model, Gaussian approximation provides better information than the large-N approximation [4]. The formalism has been successfully used to derive effective potential in $(2+1)$ dimensional Liouville model [5]. Quantum field theoretic analysis of inflation dynamics in a $(2+1)$ dimensional universe has been worked out using this method. Taking into consideration the recent applications of the formalism in $(2+1)$ dimensional theories we propose to apply the method to the most general renormalizable scalar theory in $(2+1)$ dimensions that is a Φ^6 model. We find that the effective potential expression that emerges using functional Schrödinger picture is same as that derived using CJT formalism [6, 7]. The functional Schrödinger picture formalism for quantum field theory is a generalization from ordinary quantum mechanics to infinite number of degrees of freedom that comprise a field [8]. The method is suitable for both static and time-dependent problems at zero and finite temperatures and for quantum fields far from equilibrium. It has also been shown that renormalization in this model does not pose any special difficulties for static or time-dependent problems [9–15]. Among the applications of this approach are scalar QED, quantum mechanics of inflation, quantum roll process and quantum processes in non-euclidean space-time [16].

to be superior to conventional Fock space methods for analysing detailed structural properties of a quantum field. In this method one need not choose a vacuum and normal ordering is not required. Schrödinger picture method has been extensively used in studying solitons and other collective phenomena, topological defects in gauge theories and confinement. Time dependent problems relevant to early universe studies and non-equilibrium thermal physics cannot be studied using conventional Greens's function methods which require an initial condition for the solution. In these areas time-dependent Schrödinger picture has been profitably used. Also for analysing representations of transformation groups the method is very useful [9]. One can achieve intrinsic regularization and renormalization without any reference to vacuum state and an unique representation is obtained. In this method a quantum mechanical state $|\Psi(t)\rangle$ is replaced by a functional of the c-number field $\phi(x)$

$$|\Psi(t)\rangle \rightarrow \Psi[\phi, t]. \tag{1}$$

The action of an operator can be realized as a product and that of a canonical momentum as a functional differentiation

$$\Phi(x)|\Psi(t)\rangle \rightarrow \Psi(x)\Psi(\phi, t), \tag{2}$$

$$\Pi(x)|\Psi(t)\rangle \rightarrow -i\frac{\delta}{\delta\phi(x)}\Psi(\phi, t). \tag{3}$$

The dynamical evolution of a given initial state is described by the functional Schrödinger equation. This equation can be derived using variational principles if we define a time-dependent effective action [12]

$$\Gamma = \int \mathrm{d}t\langle\Psi(t)|i\partial_t - H|\psi(t)\rangle \tag{4}$$

and impose the condition that $|\Psi(t)\rangle$ is stationary against arbitrary variations we get

$$i\frac{\partial\Psi(\phi, t)}{\partial t} = H\Psi(\phi, t) = \int_x \left[-\frac{1}{2}\frac{\delta^2}{\delta\phi^2(x)} + \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right]\Psi(\phi, t). \tag{5}$$

For applying variational method we have to prescribe a trial state. In the Gaussian approximation we assume a Gaussian trial state

$$\Psi(\phi, t) = \exp\left[-\int_{x,y}(\phi(x) - \hat{\phi}(x, t))\left[\frac{G^{-1}(x, y, t)}{4} - i\Sigma(x, y, t)\right]\right]$$
$$\times(\phi(y) - \hat{\phi}(y, t)) + i\int_x \hat{\Pi}(x, t)[\phi(x) - \hat{\phi}(x, t)]. \tag{6}$$

It can be seen that the Gaussian is centered at $\hat{\phi}$ and the width is given by G. Σ plays the role of the conjugate momentum of G and $\hat{\Pi}$ that of $\hat{\phi}$. $\hat{\phi}, \hat{\Pi}, G$ and Σ are the variational parameters as well as the expectation values:

$$\langle\phi(x)\rangle = \hat{\phi}(x, t), \tag{7}$$

$$\left\langle -i\frac{\delta}{\delta\phi(x)}\right\rangle = \Pi(\hat{x}, t), \tag{8}$$

$$\langle\phi(x)\phi(y)\rangle = \phi(\hat{x},t)\phi(\hat{y},t) + G(x,y,t), \tag{9}$$

$$\left\langle i\frac{\partial}{\partial\phi(x)}\right\rangle = \int_x \hat{\Pi}(x,t)\dot{\hat{\phi}}(x,t) + \int_{x,y}\Sigma(x,y,t)\dot{G}(y,x,t), \tag{10}$$

$$V^{(n)}(\hat{\phi}) \equiv \frac{\mathrm{d}^n V(\hat{\phi})}{\mathrm{d}\hat{\phi}^n}. \tag{11}$$

For applying the formalism to a Φ^4 model, the following expression for effective action can be written up to two loop level [10, 11]

$$\Gamma = \int \mathrm{d}t\left[\int_x \hat{\Pi}\,\dot{\hat{\phi}} - \frac{1}{2}(\nabla\hat{\phi})^2 - V(\hat{\phi}) + \int_{x,y}\Sigma\dot{G} - 2\int_{x,y,z}\Sigma G\Sigma - \int_x \frac{1}{8}G^{-1}(x,x,t)\right.$$
$$\left. - \frac{1}{2}\nabla_x^2 G(x,y,t)|_{x=y} + \frac{1}{2}V^{(2)}(\hat{\phi})G(x,x,t)\right] - \frac{1}{8}V^{(4)}(\hat{\phi})\int_x G(x,x,t)^2. \tag{12}$$

Identifying the first term as the classical action and performing variations we get

$$\frac{\delta\Gamma}{\delta\dot{\hat{\phi}}(x,t)} = 0 \rightarrow \dot{\hat{\Pi}}(x,t) = \nabla_x^2\hat{\phi}(x,t) - V^{(1)}(\hat{\phi}) - \frac{1}{2}V^{(3)}(\hat{\phi})G(x,x,t), \tag{13}$$

$$\frac{\delta\Gamma}{\delta\dot{\hat{\Pi}}(x,t)} = 0 \rightarrow \dot{\Sigma}(x,y,t) + 2\int_x \Sigma(x,z,t)\Sigma(x,y,t)$$
$$= \frac{1}{8}G^{-2}(x,y,t) + \left[\frac{1}{2}\nabla_x^2 - \frac{1}{2}V^{(2)}(\hat{\phi}) - \frac{1}{4}V^{(4)}(\hat{\phi})G(x,x,t)\right]\delta^\nu(x-y), \tag{14}$$

$$\frac{\delta\Gamma}{\delta\Sigma(x,y,t)} = 0 \rightarrow \dot{G}(x,y,t) = 2\left[\int_x G(x,z,t)\Sigma(x,y,t) + \Sigma(x,z,t)G(z,y,t)\right]. \tag{15}$$

The static effective potential can be obtained by taking $\hat{\phi}$ to be x-independent and by putting $\Sigma = 0$. By performing variation for G, a gap equation for G could be written. A slightly different but equivalent method has been used in $(2+1)$ dimensional Liouville model [5].

2. Effective action

We propose to apply the formalism to a Φ^6 model and since Φ^6 coupling effects show up only at the three loop level we write the expression up to three loops

$$\Gamma = \int \mathrm{d}t\left[\int_x \hat{\Pi}\,\dot{\hat{\phi}} - \frac{1}{2}(\nabla\hat{\phi})^2 - V(\phi) + \int_{x,y}\Sigma\dot{G} - 2\int_{x,y,z}\Sigma G\Sigma\right.$$
$$\left. - \int_x \frac{1}{8}G^{-1}(x,x,t) - \frac{1}{2}\nabla_x^2 G(x,y,t)|_{x=y} + \frac{1}{2}V^{(2)}(\hat{\phi})G(x,x,t)\right]$$
$$- \frac{1}{8}V^{(4)}(\hat{\phi})\int_x G(x,x,t)^2 - \frac{1}{16}V^{(6)}(\hat{\phi})\int_x G^3(x,x,t). \tag{16}$$

the classical action and performing variations we obtain

$$\frac{\delta\Gamma}{\delta\dot{\hat{\phi}}(x,t)} = 0 \rightarrow \dot{\hat{\Pi}}(x,t) = \nabla_x^2\hat{\phi}(x,t) - V^{(1)}(\hat{\phi}) - \frac{1}{2}V^{(3)}(\hat{\phi})G(x,x,t) - \frac{1}{8}V^{(5)}(\hat{\phi})G^2(x,x,t), \tag{17}$$

$$\frac{\delta\Gamma}{\delta\dot{\hat{\Pi}}(x,t)} = 0 \rightarrow \dot{\Sigma}(x,y,t) + 2\int_x \Sigma(x,z,t)\Sigma(x,y,t) = \frac{1}{8}G^{-2}(x,y,t) + \left[\frac{1}{2}\nabla_x^2 - \frac{1}{2}V^{(2)}(\hat{\phi}) - \frac{1}{4}V^{(4)}(\hat{\phi})G(x,x,t) - \frac{1}{4}V^{(6)}(\hat{\phi})G^2(x,x,t)\right]\delta^\nu(x-y), \tag{18}$$

$$\frac{\delta\Gamma}{\delta\Sigma(x,y,t)} = 0 \rightarrow \dot{G}(x,y,t) = 2\left[\int_x G(x,z,t)\Sigma(x,y,t) + \Sigma(x,z,t)G(z,y,t)\right]. \tag{19}$$

3. Static effective potential

It has been shown that renormalization of time-dependent equations can be done along the same lines of the renormalization of the static effective potential. We therefore evaluate the static effective potential for Φ^6 model at zero temperature. The static effective potential takes the form

$$V_{\text{eff}}(\hat{\phi}, G) = \frac{1}{2}m^2\hat{\phi}^2 + \frac{\lambda}{4!}\hat{\phi}^4 + \frac{\xi}{6!}\hat{\phi}^6 + \left(\frac{1}{2}m^2 + \frac{\lambda}{4}\hat{\phi}^2 + \frac{\xi}{48}\hat{\phi}^4\right)G(x,x) + \left(\frac{\lambda}{8} + \frac{\xi}{16}\hat{\phi}^2\right) \times G^2(x,x) + \frac{\xi}{48}G^3(x,x) + \frac{1}{8}\text{tr}\,G^{-1}(x,x) - \frac{1}{2}\nabla_x^2 G(x,x). \tag{20}$$

By performing variation with respect to G the gap equation is obtained.

$$\frac{1}{4}G^{-2}(x,y) = \left[-\delta_x^2 + m^2 + \frac{\lambda}{2}\hat{\phi}^2 + \frac{\xi}{24}\hat{\phi}^4 + \frac{\lambda}{2}G(x,y) + \frac{\xi}{4}\hat{\phi}^2G^2(x,y) + \frac{\xi}{48}G^3(x,y)\right]\delta(x-y). \tag{21}$$

Assuming translation invariance the Fourier transform of a function is defined as

$$f(x) \equiv \int \frac{\mathrm{d}^\nu k}{2\pi^\nu} e^{ikx} f(k), \tag{22}$$

$$G(x,x) = \int \frac{\mathrm{d}^\nu k}{2\pi^\nu}\frac{1}{2}\left[\left[k^2 + m^2 + \frac{\lambda}{2}\hat{\phi}^2 + \frac{\xi}{24}\hat{\phi}^4\right] \times \frac{\lambda}{2}G(x,x) + \frac{\xi}{4}\hat{\phi}^2G(x,x) + \frac{\xi}{48}G^2(x,x)\right]^{-1/2}, \tag{23}$$

$$G(x,x) = \int \frac{\mathrm{d}^\nu k}{2\pi^\nu} \frac{1}{2(k^2+M^2)^{1/2}}, \tag{24}$$

where an ansatz is fixed for G in terms of an effective mass. The effective mass M is treated here as a variational parameter which is $\hat{\phi}$ dependent. The static effective potential can be written as

$$\begin{aligned} V_{\mathrm{eff}}(\phi, M) = & \frac{1}{2}\int \frac{\mathrm{d}^\nu k}{2\pi^\nu}(k^2+M^2)^{1/2} + \left(\frac{1}{2}m_{\mathrm{B}}^2\hat{\phi}^2 + \frac{\lambda}{4!}\hat{\phi}^4 + \frac{\xi_{\mathrm{B}}}{6!}\hat{\phi}^6\right) \\ & + \frac{1}{2}\left[M^2 - m_{\mathrm{B}}^2 - \frac{\lambda_{\mathrm{B}}}{2}\hat{\phi}^2 - \frac{\xi_{\mathrm{B}}}{24}\hat{\phi}^4\right] G(x,x) \\ & + \left[\frac{\lambda}{8} + \frac{\xi}{16}\hat{\phi}^2\right] G^2(x,x) + \frac{\xi}{48}G^3(x,x). \end{aligned} \tag{25}$$

Equation (25) shows that the effective potential expression obtained here is the same as the one obtained using Gaussian effective potential approach [17]. This is only natural since when time dependence is not taken into account definitions of effective action in both the approaches coincide. It also becomes clear that the formalism is equivalent to CJT approach at zero temperature. Equations obtained here differ from those obtained in CJT formalism by a $\xi\phi^2$ term. This term does not contribute when daisy and super daisy diagrams alone are considered through Hartee–Fock approximation which is the popular approximation employed in CJT formalism. At ϕ^4 level both the approaches are exactly identical.

Considering the first and second terms alone of (25) it can be seen that one loop effective potential result is contained in the expression with the mass term replaced by the effective mass. Identity with the Gaussian effective potential results become more transparent if we make the following correspondence in notation.

$$G(x,x) \to I_0 = \int \frac{\mathrm{d}^\nu k}{2\pi^\nu} \frac{1}{2(k^2+m^2)^{1/2}}, \tag{26}$$

$$\frac{1}{2}\int \frac{\mathrm{d}^\nu k}{2\pi^\nu}(k^2+m^2)^{1/2} \to I_1, \tag{27}$$

$$M \to \Omega. \tag{28}$$

Since the effective potential is an ordinary function (not a functional) stationary requirements with respect to ϕ and M^2 is obtained by ordinary differentiation

$$\frac{\partial V}{\partial \phi} = \phi\left[m_{\mathrm{B}}^2 + \frac{\lambda_{\mathrm{B}}}{6}\phi^2 \frac{\xi_{\mathrm{B}}}{120}\phi^4 + \frac{\lambda_{\mathrm{B}}}{2}G(x,x) + \frac{\xi_{\mathrm{B}}}{12}\Phi^2 G(x,x) + \frac{\xi}{8}G(x,x)G(x,x)\right] = 0 \tag{29}$$

$$\begin{aligned} \frac{\partial V}{\partial M^2} = -\frac{1}{2}\Bigg[M^2 - m_{\mathrm{B}}^2 - \frac{\lambda_{\mathrm{B}}}{2}\phi^2 - \frac{\xi_{\mathrm{B}}}{24}\phi^4 - \frac{\lambda_{\mathrm{B}}}{2}G(x,x) - \frac{\xi_{\mathrm{B}}}{4}\hat{\phi}^2 G(x,x) \\ - \frac{\xi_{\mathrm{B}}}{8}G(x,x)G(x,x)\Bigg]\frac{\partial G(x,x)}{\partial M^2} = 0. \end{aligned} \tag{30}$$

Conventional effective potential is defined at the solution of (30). The effective mass is

given by

$$M^2(\phi) = \left[m_B^2 + \frac{\lambda_B}{2}\phi^2 + \frac{\xi_B}{24}\phi^4 + \frac{\lambda_B}{2}G(x,x) + \frac{\xi_B}{8}G(x,x)G(x,x)\right]. \tag{31}$$

Required expression for the effective potential is obtained by replacing the effective mass M by $M(\phi)$ in eq. (25). Equation (25) shows certain very important peculiarities of Φ^6 and Φ^4 theories relevant at zero temperature.

$$V'(\phi) = \phi\left[M^2(\phi) - \left(\frac{\lambda_B}{3}\phi^2\right)\right]. \tag{32}$$

For Φ^4 theory $M^2(\phi)$ is intrinsically positive. Hence if $\lambda_B < 0$ only solution to the above equation is $\phi = 0$ (or potential is unbounded from below). That is for negative λ_B non-zero turning points do not exist. In the case of Φ^6 theory

$$V'(\phi) = \phi\left[M^2(\phi) - \left(\frac{\lambda_B}{3}\phi^2 + \frac{\xi_B}{30}\phi^4\right)\right] - \frac{\xi}{4}\left[1 - \frac{\phi^2}{3}\right]G^2 \tag{33}$$

non-zero turning points are possible also for $\lambda_B < 0$. Φ^6 theory in Hartee–Fock approximation requires up to three loops for obtaining the effects of ξ_B coupling. We have four parts for the effective potential

$$V_\beta(\phi, M(\phi)) = V^0 + V^1 + V^2 + V^3, \tag{34}$$

$$V^0 = \left[\frac{1}{2}m_B^2\phi^2 + \frac{\lambda_B}{4!}\phi^4 + \frac{\xi_B}{6!}\phi^6\right], \tag{35}$$

$$V^1 = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\ln[k^2 + M^2(\phi)], \tag{36}$$

$$V^2 = -\frac{\lambda_B}{8}G(x,x)G(x,x) - \frac{xi}{16}(\phi)^2G(x,x)G(x,x), \tag{37}$$

$$V^3 = -\frac{\xi_B}{24}G(x,x)G(x,x)G(x,x), \tag{38}$$

which are obtained by substituting $M(\phi)$. Effective potential for both Φ^4 and Φ^6 theories can be obtained from this equation. In the following we do not consider the $\hat{\phi}^2$ dependent term to establish the equivalence with the Hartee–Fock approximation commonly used in the CJT formalism [6].

4. Renormalization

The effective mass term $M(\phi)$ defining the effective potential is divergent mainly due to the presence of $G(x,x)$. Following re-normalization prescription is employed in $(2+1)$ dimensions to regularize $M(\phi)$. Define

$$G(M(\phi)) \equiv -\frac{M(\phi)}{4\pi} \tag{39}$$

with

$$E \equiv [k^2 + M^2(\phi)]^{1/2}. \tag{40}$$

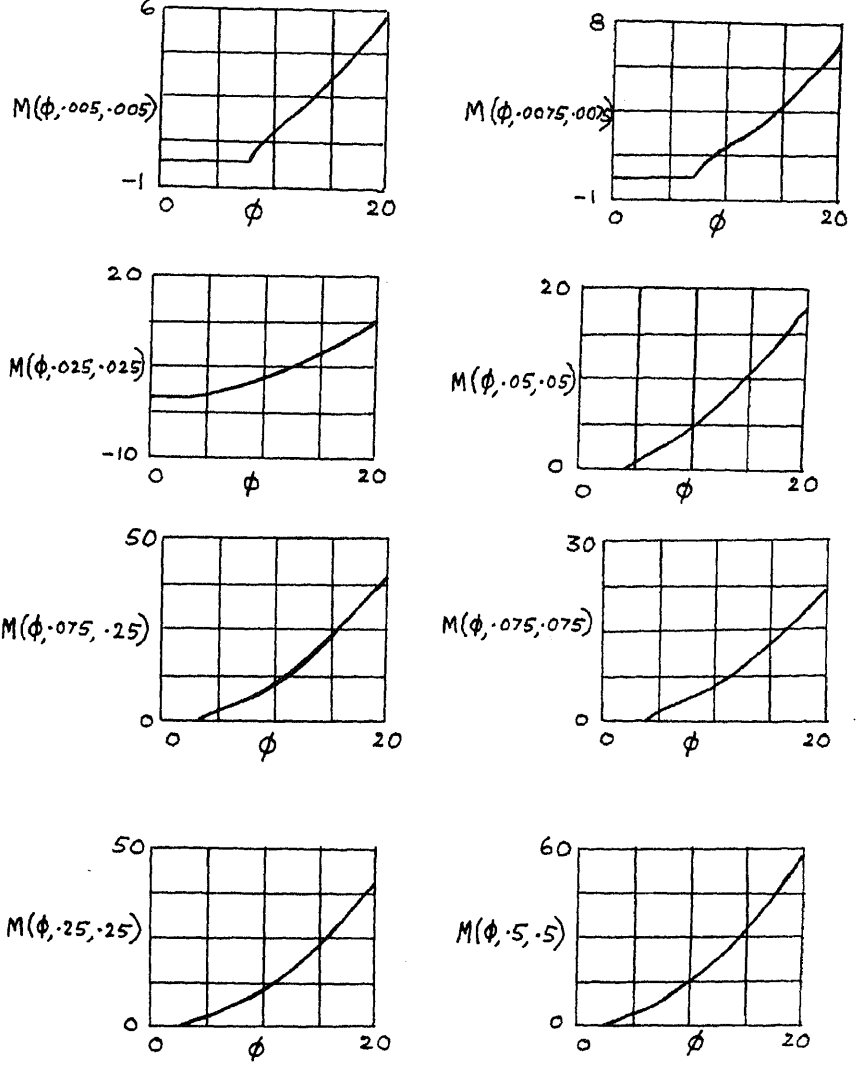


Figure 1. Effective mass for various couplings-Φ^6 theory.

In $(2+1)$ dimensions coupling constant renormalization is not required. Define

$$m_r^2 \equiv m_B^2 + \frac{1}{2}\lambda I_1 + \frac{\xi'}{4} I_1 + \frac{\xi}{8} I_1^2, \tag{41}$$

$$I_1 \equiv \int \frac{d^2k}{(2\pi)^2} \frac{1}{2k} = \lim_{\Lambda\to\infty} \left(\frac{\Lambda}{4\pi}\right). \tag{42}$$

$\xi' \equiv \xi G(M(\phi))$,

$$G(x,x) = \int \frac{d^2k}{(2\pi)^2} \frac{1}{2E}. \tag{43}$$

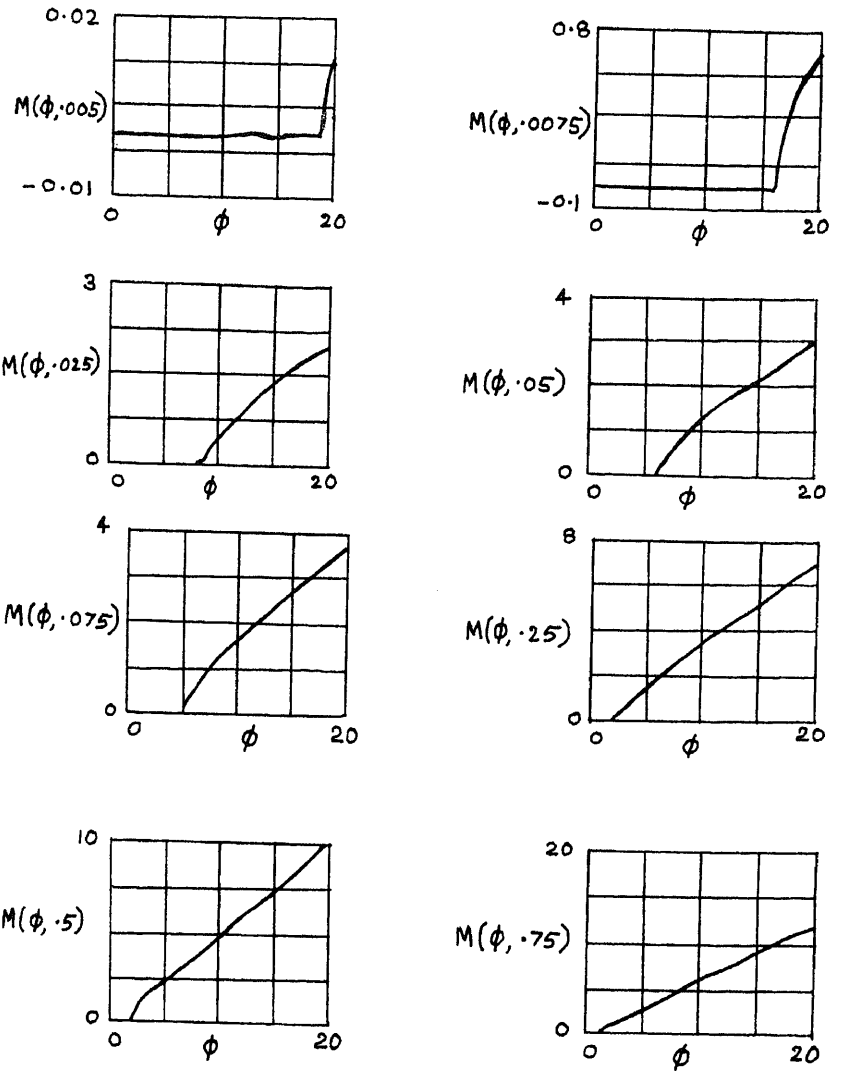


Figure 2. Effective mass for various couplings-Φ^4 theory.

By actual evaluation, introducing a cut-off parameter Λ

$$G(x,x) = G(M(\phi)) + I_1. \tag{44}$$

Equation (34) shows that $G(M(\phi))$ is the finite part of the vacuum propagator. A finite expression for $M(\Phi)$ is obtained by expressing it in terms of the renormalized parameters:

$$M^2(\phi) = -m_r^2 + \frac{\lambda}{2}\phi^2 + \frac{\xi}{24}\phi^4 + \frac{\lambda}{2}G(M(\phi)) + \frac{\xi}{8}G(M(\phi))G(M(\phi)). \tag{45}$$

Second derivative of the tree level potential is defined as m_{tree} (tree level mass)

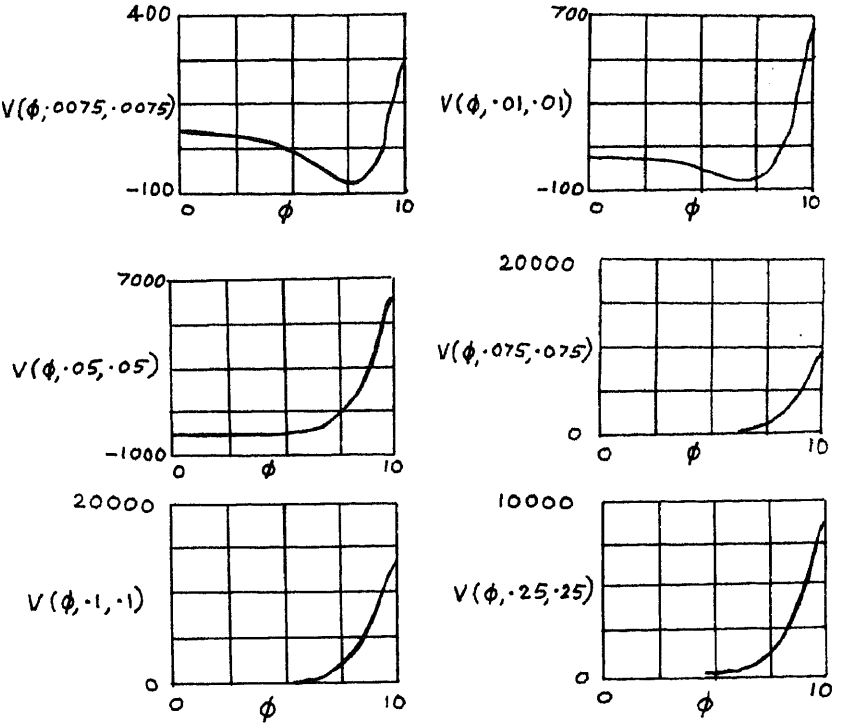


Figure 3. Effective potential for various couplings-Φ^6 theory.

$$M^2(\phi) = m^2_{\text{tree}}(\phi) + \frac{\lambda}{2} G(M(\phi)) + \frac{\xi}{8} G(M(\phi)) G(M(\phi)), \tag{46}$$

$$V^1(M(\phi)) = -\frac{M^3}{6\pi} + \frac{\Lambda^3}{6\pi}. \tag{47}$$

The last term is cut-off dependent. Cancellation of this divergence is obtained by combining V^0, V^2 and V^3:

$$V(M) = -\frac{M^3}{6\pi} + \frac{M^4}{2\lambda} - \frac{1}{2} M^2 G(M) - \frac{M^2}{24\eta} G^2(M) - F(\phi) \tag{48}$$

with

$$\eta^{-1} = \frac{\xi}{\lambda} \text{ and } F(\phi) = \left[\frac{m^2}{24\eta} - \frac{\lambda}{12} - \frac{29}{6!}\xi\phi^2\right]\phi^4. \tag{49}$$

$\xi = 0$ reproduces the result of Camellia and Pi at zero temperature. Using unrenormalized gap equation we combine V^0, V^2 and V^3 and writing them in terms of renormalized parameters

$$V^0 + V^2 + V^3 = \frac{\lambda_r}{8}\left[\phi^2 - \frac{2m_r^2}{\lambda}\right]^2 - \frac{\eta}{24} m_r^2 \phi^4 - \frac{\lambda_r}{8} G^2(M) - \frac{\xi_r}{48} G^3(M) - F(\phi). \tag{50}$$

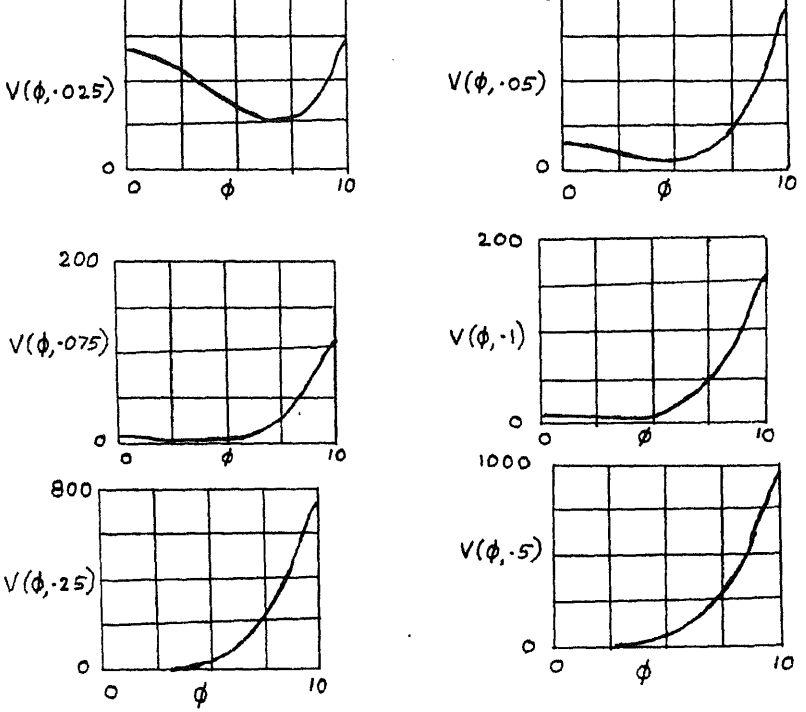


Figure 4. Effective potential for various couplings-Φ^4 theory.

In the case of $(2+1)$ dimensional Φ^4 theory $F(\phi) = \lambda/12$ which is finite. Thus unlike $(3+1)$ dimensional Φ^4 theory the effective potential do not contain any unrenormalized parameters. But in the case of Φ^6 theory $F(\phi)$ contains m which is an unrenormalized mass parameter. But here we can make $F(\phi) = 0$ by adjusting the parameters suitably and make the unrenormalized parameters vanish.

5. Conclusions

Static effective potential for Φ^6 theory is obtained using the functional Schrödinger picture formalism for calculating the effective action. Regarding the presence of turning points it is shown that for $(2+1)$ dimensional Φ^6 theory turning points can exist both for positive and negative λ. This indicates the possibility of bound states in the case of Φ^6 theories. The equations obtained for the effective potential and effective mass are found to be identical with the more popular Gaussian effective potential. Apart from a difference arising due to the use of Hartee–Fock approximation the formalism is identical to the CJT formalism.

Graphical analysis of the effective mass and effective potential is performed by numerically solving the equations for effective mass and effective potential. Graphs plotted compared with earlier works show close similarity with CJT formalism.

Application of the formalism to finite temperature situation is presently under investigation.

Comparing the graphs obtained for Φ^4 and Φ^6 theories. (figures 1 and 2) we can see the similarity in behaviour. Value of the effective mass for Φ^4 theory is smaller as expected. Straight portion of the graph indicating imaginary value of effective mass is more pronounced in the case of Φ^4 theory. The straight portion of the graph is more pronounced also for weak couplings. Comparison between effective potential values can be seen from figures 3 and 4.

References

[1] H J Lee, K Na and J H Yee, *Phys. Rev.* **D51**, 3125 (1995)
[2] C Keifer, *Phys. Rev.* **D45**, 2044 (1992)
[3] S K Kim, W Namgung, K S Soha and J H Yee, *Phys. Rev.* **D41**, 1209 (1990)
[4] S Hyun, G H Lee and Y H Yee, *Phys. Rev.* **D50**, 6542 (1994)
[5] B K Majumdar and R K Roychoudhari, *Int. J. Theor. Phys* **21**, 1389 (1992)
[6] G A Camelia and S Y Pi, *Phys. Rev.* **D47**, 2356 (1993)
[7] J M Cornwall, R Jackiw and E Tomboulis, *Phys. Rev.* **D10**, 2428 (1974)
[8] O Eboli, R Jackiw and S Y Pi, *Phys. Rev.* **D37**, 3557 (1988)
[9] R Jackiw, in *Conformal field theory anomalies and super strings* edited by B Baaquie, C Chew, C Oh and K K Phua (World Scientific, Singapore, 1988)
[10] F Cooper, S Y Pi and P N Stancioff, *Phys. Rev.* **D34**, 3831 (1986)
[11] S Y Pi and M Samiullah, *Phys. Rev.* **D36**, 3128 (1987)
[12] R Jackiw and A Kerman, *Phys. Lett.* **A71**, 158 (1979)
[13] K Symanzik, *Nucl. Phys.* **B190**, 1 (1983)
[14] R Floreamini and R Jackiw, *Phys. Rev.* **D37**, 2206 (1988)
[15] A H Guth and S Y Pi, *Phys. Rev.* **D32**, 1899 (1985)
[16] S K Kim, J Yang, K S Soha and J H Yee, *Phys. Rev.* **D40**, 2647 (1989)
[17] G A Hajj and P M Stevenson, *Phys. Rev.* **D37**, 413 (1988)
M Samiullah, O Eboll and S Y Pi, *Phys. Rev.* **D44**, 2335 (1991)

Motion of test particles around domain walls

SUBENOY CHAKRABORTY and LALIT BISWAS*
Department of Mathematics, Jadavpur University, Calcutta 700 032, India
*Regional Meteorological Centre, India Meteorological Department, Govt. of India, Calcutta 700 027, India

MS received 3 February 1997; revised 4 August 1997

Abstract. We present a detailed analysis of the motion of test particles around domain walls. The study of the trajectories of the test particles has been done using the Hamilton–Jacobi formalism. In most of the cases we show that the particles can not be trapped by the walls.

Keywords. Domain walls; test particles; trajectories; Hamilton–Jacobi formalism.

PACS Nos 04.20; 98.80

1. Introduction

The explanation of the formation of large scale structure of the Universe is one of the basic problem of cosmology even today. In literature, the widely used mechanisms for this structure formation are the gravitational perturbations generated by topological defects namely, domain walls, cosmic strings, monopoles and their hybrids [1]. It is generally assumed that these topological defects are formed at the very early stages of the evolution of the Universe when the Universe exhibits a series of phase transitions [2]. In particular, the appearance of domain walls is associated with the breaking of a discrete symmetry [3] i.e. the vacuum manifold consists of several disconnected components.

From the cosmological standpoint, domain walls have become important after a proposal for a new scenario of galaxy formation by Hill, Schramn and Fry [4]. According to them, the formation of galaxies are due to domain walls produced during a phase transition after the time of recombination of matter and radiation [5]. In fact, the phase transition occurs by the breaking of a discrete symmetry of a weakly coupled scalar field of pseudo-gold stone bosons ([3, 5]).

In this paper, we study the motion of test particles in the gravitational field of static and non-static planar domain walls and of plane symmetric domain walls using Hamilton–Jacobi (H–J) formalism and examine whether bound orbits are possible or not. This study will be interesting in the relevance of structure formation because it gives us some idea about the behaviour of the particles (created at the early Universe) in the gravitational field of the domain walls.

2. Plane symmetric domain walls

2.1 *Thick walls*

The line element describing plane symmetric thick domain walls is given by [5]

$$dS^2 = \frac{\text{sech}^{2n}(z)}{L^2}[dt^2 - n^2dz^2 - e^{2t}(dx^2 + dy^2)], \tag{1}$$

where L and n are constants. The horizon in all directions are of the order of $1/L$ and n is involved in the relation between pressure perpendicular to the wall and the energy density.

We consider a relativistic particle with mass m moving in the field of the plane symmetric thick domain wall (1). The H–J equation has the expression

$$L^2\cosh^{2n}(z)[(\partial S/\partial t)^2 - 1/n^2(\partial S/\partial z)^2 - e^{-2t}\{(\partial S/\partial x)^2 + (\partial S/\partial y)^2\}] + m^2 = 0. \tag{2}$$

As the metric (1) is independent of x and y co-ordinates, a natural choice for H–J function in separable form will be

$$S(x, y, z, t) = p_x \cdot x + p_y \cdot y + S_1(z) + S_2(t). \tag{3}$$

Here, the constants p_x and p_y can be termed as the momentum of the particle in the plane of the wall. If we substitute the ansatz (3) for S in the H–J equation (2) then the expression (in integral form) for the unknown functions S_1 and S_2 are

$$S_1(z) = \epsilon n \cdot \int \left[-E^2 + \frac{m^2}{L^2}\,\text{sech}^{2n}(z)\right]^{1/2} dz, \tag{4a}$$

$$S_2(t) = \epsilon \cdot \int [-E^2 + p^2 e^{-2t}]^{1/2} dt, \tag{4b}$$

where $P^2 = p_x^2 + p_y^2$, E is a constant and $\epsilon = \pm 1$ stands for the sign changing whenever z (or t) passes through a zero of the integrand in (4a) or (4b) [6, 7].

In the H–J formalism, the path of the particle is characterized by [7, 8]

$$\frac{\partial S}{\partial p_x} = \text{constant}, \quad \frac{\partial S}{\partial p_y} = \text{constant and } \frac{\partial S}{\partial E} = \text{constant}.$$

Thus we get

$$x = \epsilon \int \frac{p_x \cdot e^{-2t}}{(p^2 e^{-2t} - E^2)^{1/2}}\, dt, \tag{5}$$

$$y = \epsilon \int \frac{p_y \cdot e^{-2t}}{(p^2 e^{-2t} - E^2)^{1/2}}\, dt, \tag{6}$$

and

$$\int \frac{dz}{[(m^2/L^2)\text{sech}^{2n}(z) - E^2]^{1/2}} = \epsilon \int \frac{dt}{n} \cdot \frac{1}{\sqrt{P^2 \cdot e^{-2t} - E^2}}, \tag{7}$$

where we have taken the constants to be zero without any loss of generality. From (7) the velocity of the particle transverse to the wall is

$$\frac{dz}{dt} = \frac{1}{n} \frac{[(m^2/L^2)\text{sech}^{2n}(z) - E^2]^{1/2}}{[P^2 e^{-2t} - E^2]^{1/2}}. \tag{8}$$

The turning points of the trajectory are given by $dz/dr = 0$ and as a consequence the potential curves are characterized by

$$\frac{E}{m} = \frac{1}{L} \text{sech}^n(z). \tag{9}$$

As there are no real extrema (except for $z = 0$) for the potential curve, the particles can not be trapped by plane symmetric thick domain walls. This conclusion is also supported by Goetz [5] that wall described by metric (1) exerts repulsive gravitational field on the matter around it.

2.2 *Thin wall*

We now consider a plane thin domain wall located at $z = 0$. The metric ansatz describing the gravitational field of the Wall is [9]

$$dS^2 = e^{-4\pi\sigma|z|}(dt^2 - dz^2) - e^{4\pi\sigma(t-|z|)}(dr^2 + r^2 d\phi^2), \tag{10}$$

where σ is the surface tension of the wall.

As the metric coefficients are independent of ϕ co-ordinate, the solution of the H–J equation

$$\begin{aligned} &e^{4\pi\sigma|z|}[(\partial S/\partial t)^2 - (\partial S/\partial z)^2] \\ &\quad = e^{-4\pi\sigma(t-|z|)}[(\partial S/\partial r)^2 + (1/r)^2(\partial S/\partial\phi)^2] + m^2 = 0, \end{aligned} \tag{11}$$

can be expressed in the following separable form

$$S(t, z, r, \phi) = S_1(t) + S_2(z) + S_3(r) + J \cdot \phi. \tag{12}$$

The constant J can be identified as the angular momentum of the test particle and the unknown functions have the integral expressions

$$\begin{aligned} S_1 &= \epsilon \int (p_1^2 \cdot e^{-4\pi\sigma t} - E_1^2)^{1/2} dt, \\ S_2 &= \epsilon \int (m^2 e^{-4\pi\sigma|z|} - E_1^2)^{1/2} dz, \\ S_3 &= \epsilon \int (p_1^2 \cdot \tau J^2/r^2)^{1/2} dr, \end{aligned} \tag{13}$$

where E_1 and p_1 are separable constants. Now, the equation of the trajectories can be written as

$$\int (p_1^2 \cdot e^{-4\pi\sigma t} - E_1^2)^{-1/2} dt = \epsilon \int (m^2 \cdot e^{-4\pi\sigma|z|} - E_1^2)^{-1/2} dz, \tag{14}$$

$$\int (p_1^2 - J^2/r^2) \qquad \mathrm{d}r = \epsilon \int \frac{}{(p_1^2 \mathrm{e}^{-4\pi\sigma t} - E_1^2)^{1/2}}, \tag{15}$$

$$\phi = \epsilon \int \frac{J}{r^2} \cdot \frac{\mathrm{d}r}{\sqrt{p_1^2 - J^2/r^2}}. \tag{16}$$

Hence from (14), the expression for transverse velocity is

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \left\{ \frac{(m^2 \cdot \mathrm{e}^{-4\pi\sigma|z|} - E_1^2)}{(p_1^2 \mathrm{e}^{-4\pi\sigma t} - E_1^2)} \right\}^{1/2}. \tag{17}$$

The turning points are given by

$$\frac{E_1}{m} = \mathrm{e}^{-2\pi\sigma|z|}. \tag{18}$$

As the expression on the right hand side is monotonic decreasing for $z > 0$ and $z < 0$ so bound states are not possible and particle experiences a repulsive force near the wall.

Therefore, particles cannot be trapped by plane domain wall whether it is thin or thick.

3. Planar domain wall

We now consider static planar wall with metric ansatz of the form [10]

$$\mathrm{d}S^2 = f^2(|x|)\mathrm{d}t^2 - \frac{1}{f^2(|x|)}\,\mathrm{d}x^2 - g^2(|x|)(\mathrm{d}y^2 + \mathrm{d}z^2), \tag{19}$$

where

$$f^2(x) = \frac{C}{x - x_0} - \frac{\Lambda}{3}(x - x_0)^2, \quad g^2(x) = K^2(x - x_0)^2,$$

and C, K and x_0 are constants satisfying the inequality

$$C + \frac{\Lambda}{3} x_0^3 < 0.$$

In this case the H–J equation has the expression

$$\frac{1}{f^2(|x|)}\left(\frac{\partial S}{\partial t}\right)^2 - f^2(|x|)\left(\frac{\partial S}{\partial x}\right)^2 - \frac{1}{g^2(|x|)}\left\{\left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right\} + m^2 = 0. \tag{20}$$

The solution can be written as

$$S = -E_s t + S_1(x) + p_y \cdot Y + p_z \cdot Z, \tag{21}$$

where E_s, p_y and p_z are constants and $S_1(x)$ has the expression (in integral form)

$$S_1(x) = \epsilon \int \frac{1}{f(x)} \left[\frac{E_s^2}{f^2} = \frac{p^2}{g^2} + m^2 \right]^{1/2} \mathrm{d}x. \tag{22}$$

The trajectory of the particle is characterized by

$$t = \epsilon \int \left(\frac{E}{f^3}\right) \cdot \frac{\mathrm{d}x}{[E_s^2/f^2 - p^2/g^2 + m^2]^{1/2}}, \tag{23}$$

$$y = \epsilon \int \left(\frac{p_y}{fg^2}\right) \cdot \frac{\mathrm{d}x}{[E_s^2/f^2 - p^2/g^2 + m^2]^{1/2}}, \tag{24}$$

$$z = \epsilon \int \left(\frac{p_z}{fg^2}\right) \cdot \frac{\mathrm{d}x}{[E_s^2/f^2 - p^2/g^2 + m^2]^{1/2}}. \tag{25}$$

Hence the expression for transverse velocity is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \left(\frac{f^3}{E_s}\right) \cdot \left(\frac{E_s^2}{f^2} - \frac{p^2}{g^2} + m^2\right)^{1/2}. \tag{26}$$

The turning points are given by

$$\frac{E_s}{m} = \left\{f^2 \cdot \left(\frac{p^2}{m^2 g^2} - 1\right)\right\}^{1/2} \tag{27}$$

which determines the potential curves.

As the radicand in the above expression may have real extremals, depending on the nature of the arbitrary constants in the metric co-efficients, the trajectory of the test particle is bounded i.e. particle can be trapped by Walls.

Finally, we note that $x = x_0 + (3C/\Lambda)^{1/3} = x$, corresponds to an event horizon where spatial and temporal co-ordinates interchange their roles and the regions $x < \bar{x}$ and $x > \bar{x}$ cannot influence each other.

Therefore, we conclude that domain walls are not always exert repulsive forces. There are situations where there is a force of attraction and particles are trapped by the walls.

Acknowledgement

The authors wish to thank the members of the Relativity Centre, Department of Physics, Jadavpur University, Calcutta for valuable discussions.

References

[1] T W B Kibble, *J. Phys.* **A9**, 1387 (1976)
[2] A Vilenkin, *Phys. Rep.* **121**, 263 (1985)
A D Linde, *Rep. Prog. Phys.* **42**, 389 (1979)
[3] A Wang, *Mod. Phys. Lett.* **A9**, 3605 (1994)
[4] C T Hill, D N Schrann and J N Fry, *Comm. Nucl. Part. Phys.* **19**, 25 (1989)
[5] G Goetz, *J. Math. Phys.* **31**, 2683 (1990)
[6] J L Synge and B A Griffith, *Principles of Mechanics* (McGraw-Hill Book Company, 1970)
[7] S Chakraborty, *Gen. Relativ. Gravit.* **28**,.1115 (1996)
[8] S Chakraborty and L Biswas, *Class Quant. Gravit.* **13**, 2153 (1996)
[9] P S Letelier, *Class. Quant. Gravit.* **6**, L207 (1989)
A Vilenkin and E P S Shellard, *Cosmic strings and other topological defects* (Cambridge University Press, 1994)
[10] B Linet, *Int. J. Theor. Phys.* **24**, 1159 (1985)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 609–615

Einstein pseudotensor and total energy of the universe

N BANERJEE and S SEN
Relativity and Cosmology Research Centre, Department of Physics, Jadavpur University, Calcutta 700 032, India

MS received 17 September 1996; revised 3 May 1997

Abstract. The total energy of the universe has been calculated assuming that it is the sum of the contributions from the matter part and gravitational part. The calculations involve the use of Einstein pseudotensor. Calculations have been carried out in some specific examples of spacetime geometries. In some cases the total energy is indeed zero confirming previous results but in other cases the total energy is nonzero. So Rosen's idea that the pseudotensorial calculations will lead to the result that the total energy of the universe is zero, is very much model dependent.

Keywords. Energy of the universe; pseudotensors.

PACS No. 04.20

1. Introduction

It is well known that quite a few descriptions of the origin of the universe led to the conjecture that the total energy of the universe should be zero. Investigations of Albrow [1] and Tyron [2] assume that the universe might have arisen as a quantum fluctuation of the vacuum. After the success of the inflationary cosmological models in solving some outstanding problems in standard cosmology, the idea of vacuum fluctuations has been developed further. Inspired by these results, there have been attempts to show that the total energy of the universe is zero even from a purely classical point of view. The total energy of the universe, however, is not really zero if it is calculated on the basis of the matter part alone. Recently, there have been several attempts to include the gravitational energy along with the matter part so that the total energy is zero. We refer to the works of Cooperstock [3] and Johri *et al* [4] for brief but excellent reviews.

Cooperstock [3] expressed the covariant conservation law

$$T^{\mu\nu}_{;\nu} = 0 \tag{1}$$

(where $\mu, \nu = 0, 1, 2, 3$) in the form of an ordinary divergence

$$\left[\sqrt{-g}\left(T^0_0 - \frac{3}{8\pi}\left(\frac{\dot{R}^2}{R^2} + \frac{k^2}{R^2}\right)\right)\right]_{,0} = 0, \tag{2}$$

where the spacetime is described by conformal Friedman–Robertson–Walker (FRW) metric

$$ds^2 = R^2(t)\left[dr^2 - \frac{dr^2}{(1-kr^2)} - r^2 d\Omega^2\right]. \tag{3}$$

He arrived at (2) by making use of some calculations involving Killing vectors and concluded that the total density of the universe is zero.

Rosen [5], on the other hand, used the Einstein pseudotensor formalism and showed that the total energy of the closed FRW universe is zero. Very recently Johri *et al* [4] also considered the pseudotensorial calculations to include the contribution from the gravitational part and arrived at the result that the total energy of the spatially closed FRW universe is zero and so also the total energy enclosed within any finite volume of the spatially flat FRW universe.

In this connection, it deserves mention that Raychaudhuri and Banerji [6] investigated the possibility of calculation of the energy of the universe using a different approach. They considered a fluid sphere embedded in an otherwise empty closed FRW universe and found that when the radius of the sphere goes to infinity, the mass of the sphere tends to zero. The advantage of this work is that it does not involve any pseudotensor and thus the results obtained are independent of the coordinate frame. But the problem of this approach is that when the radius tends to infinity, nothing can be said about the exterior spacetime in which the fluid sphere is embedded.

In the present work, we examine the calculation of the total energy on the basis of pseudotensor a bit critically. Pseudotensorial calculations are always dangerous as they are very much coordinate dependent and thus may lead to ambiguous results except, however, in asymptotically flat spacetimes if one is using quasi-Minkowskian coordinates [7]. Rosen's idea was that the Einstein pseudotensor gives the gravitational energy and together with the matter part yields the result that the total energy of the universe is zero. In what follows, we test Rosen's idea against some examples. Some anisotropic cosmological models are considered, in which the nonzero components of the pseudotensor are used to calculate the total energy of the universe following Rosen's idea. It is interesting to note that although in some cases the total energy indeed comes out to be zero, in some other cases the energy integral is in fact nonzero. In § 2 we discuss specific examples and in § 3 we discuss the results.

2. Calculation of the total energy

In order to write down the covariant conservation equation (1) in the form of an ordinary divergence equation, Einstein introduced a pseudotensor [8] t_ν^μ defined by the relation

$$t_\nu^\mu = \frac{1}{16\pi}\left[\mathcal{L}\,\delta_\nu^\mu - g_{\alpha\beta,\nu}\frac{\partial \mathcal{L}}{\partial g_{\alpha\beta,\mu}}\right], \tag{4}$$

where

$$\mathcal{L} = g^{\alpha\beta}(\Gamma^\nu_{\alpha\mu}\Gamma^\mu_{\beta\nu} - \Gamma^\mu_{\mu\nu}\Gamma^\nu_{\alpha\beta}). \tag{5}$$

With this definition, Einstein wrote the ordinary divergence equation in the form

$$[\sqrt{-g}(T_\nu^\mu + t_\nu^\mu)]_{,\mu} = 0. \tag{6}$$

In fact, the quantity $\sqrt{-g}(T_\nu^\mu + t_\nu^\mu)$ can be expressed as an ordinary divergence of a quantity $\mathcal{H}_\nu^{\mu\alpha}$,

$$\sqrt{-g}(T_\nu^\mu + t_\nu^\mu) = \frac{1}{16\pi}\mathcal{H}_\nu^{\mu\alpha}, \alpha, \tag{7}$$

where

$$\mathcal{H}_\nu^{\mu\alpha} = \frac{1}{\sqrt{-g}} g_{\nu\beta}[-g(g^{\mu\beta}g^{\alpha\sigma} - g^{\alpha\beta}g^{\mu\sigma})]_{,\sigma}. \tag{8}$$

It should, however, be noted that as $\mathcal{L}$ is not a scalar, the quantities t_ν^μ do not form a tensor. For details of the calculations we refer to Rosen [5] and references therein.

Rosen argued that t_ν^μ accounts for the contribution from the gravitational energy towards the total energy density and the integral of the quantity $\sqrt{-g}(T_0^0 + t_0^0)$ over all space has the significance of the total energy of the universe. Thus E, the total energy of the universe, is given by

$$E = \int \sqrt{-g}(T_0^0 + t_0^0)\mathrm{d}^3x, \tag{9}$$

where the integral extends over the whole spatial volume.

With the spacetime geometry described by the closed FRW metric of the form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{R^2(t)}{1 + r^2/4}(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2), \tag{10}$$

where $r^2 = x^2 + y^2 + z^2$, Rosen proved that E, given by (9), is indeed equal to zero.

A few examples from the anisotropic Bianchi type line elements has been picked up for similar kind of investigation.

I. *Bianchi type I*

The line element in this case is

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{e}^{2l}\mathrm{d}x^2 - \mathrm{e}^{2m}\mathrm{d}y^2 - \mathrm{e}^{2n}\mathrm{d}z^2, \tag{11}$$

where l, m, n are functions of t alone. From (7), we get

$$\sqrt{-g}(T_0^0 + t_0^0) = \frac{1}{16\pi}\mathcal{H}_0^{0\alpha}, \alpha, \tag{12}$$

and by using the metric given by (11) one can obtain from eq. (8),

$$\mathcal{H}_0^{0\alpha} = 0 \tag{13}$$

for all α. Therefore, in this case

$$\sqrt{-g}(T_0^0 + t_0^0) = \frac{1}{16\pi}\mathcal{H}_0^{0\alpha}, \alpha = 0$$

and hence the total energy of the universe

$$E = \int \sqrt{-g}(T_0^0 + t_0^0)\mathrm{d}^3x = 0 \tag{14}$$

for any finite volume.

As Bianchi type I line element reduces to the spatially flat FRW metric for $l = m = n$, (14) confirms the results obtained by Johri *et al* [4] that the energy contained in any finite volume of flat FRW universe is zero.

II. *Bianchi type II*

Bianchi type II line element can be written in the form

$$ds^2 = dt^2 - S^2(t)(dx + \tfrac{1}{2}z^2 dy)^2 - R^2(t)(z^2 dy^2 + dz^2) \tag{15}$$

which once again yields from (8) that $\mathcal{H}_0^{0\alpha} = 0$ for all α, so that in this case also

$$\sqrt{-g}(T_0^0 + t_0^0) = \frac{1}{16\pi}\mathcal{H}_{0,\alpha}^{0\alpha} = 0,$$

and hence

$$E = \int \sqrt{-g}(T_0^0 + t_0^0)d^3x \tag{16}$$

for any finite spatial volume.

III. *Bianchi type III*

The spacetime is described by the metric

$$ds^2 = dt^2 - a_1 e^{-2mz}dx^2 - a_2 dy^2 - a_3 dz^2, \tag{17}$$

where a_1, a_2, a_3 are functions of time while m is a constant. The only nonvanishing component of $\mathcal{H}_0^{0\alpha}$ is

$$\mathcal{H}_0^{03} = 2m\sqrt{\left(\frac{a_1 a_2}{a_3}\right)}e^{-mz}. \tag{18}$$

Thus,

$$\begin{aligned}\sqrt{-g}(T_0^0 + t_0^0) &= \frac{1}{16\pi}\mathcal{H}_{0,\alpha}^{0\alpha} \\ &= \frac{1}{16\pi}\mathcal{H}_{0,3}^{03} \\ &= -\frac{m^2}{8\pi}\sqrt{\left(\frac{a_1 a_2}{a_3}\right)}e^{-mz}\end{aligned} \tag{19}$$

and

$$\begin{aligned}E &= \int \sqrt{-g}(T_0^0 + t_0^0)d^3x \\ &= -\frac{m^2}{8\pi}\sqrt{\left(\frac{a_1 a_2}{a_3}\right)}\int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy \int_{z_1}^{z_2} e^{-mz}dz,\end{aligned} \tag{20}$$

which is readily integrable and evidently yields $E \neq 0$ for a finite spatial volume. Furthermore, E diverges to infinitely large value if the limits of integration are extended to infinity (x, y, z going from $-\infty$ to $+\infty$).

IV. *Bianchi type V*

In this case the general metric is

$$ds^2 = dt^2 - e^{2z}(b_1^2 dx^2 + b_2^2 dy^2) - b_3^2 dz^2 - 2b_{13} dx\, dz, \tag{21}$$

e b_1, b_2, b_3 and b_{13} are all functions of time alone. In order to make the calculations ler, we choose b_{13} to be zero, so that the metric becomes diagonal as

$$ds^2 = dt^2 - e^{2z}(b_1^2 dx^2 + b_2^2 dy^2) - b_3^2 dz^2. \tag{22}$$

e relevant nonvanishing component of $\mathcal{H}_0^{0\alpha}$ is

$$\mathcal{H}_0^{03} = -\frac{4b_1 b_2}{b_3} e^{2z} \tag{23}$$

o,

$$\begin{aligned}\sqrt{-g}(T_0^0 + t_0^0) &= \frac{1}{16\pi} \mathcal{H}_{0,\alpha}^{0\alpha} \\ &= \frac{1}{16\pi} \mathcal{H}_{0,3}^{03} \\ &= -\frac{e^{2z}}{2\pi} \frac{b_1 b_2}{b_3}.\end{aligned} \tag{24}$$

otal energy E is therefore given by

$$\begin{aligned}E &= \int \sqrt{-g}(T_0^0 + t_0^0) d^3x \\ &= -\frac{b_1 b_2}{2\pi b_3} \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy \int_{z_1}^{z_2} e^{2z} dz.\end{aligned} \tag{25}$$

is also readily integrable and again, for a finite spatial volume, does not vanish. Here E attains an infinitely large value if the limits of integration are extended to infinity. should note that Bianchi type V model is the anisotropic generalization of an open model.

anchi type VI

$$ds^2 = dt^2 - c_1 e^{-nz} dx^2 - c_2 e^{z} dy^2 - c_3 dz^2, \tag{26}$$

e c_1, c_2, c_3 are functions of time alone and n is a constant. The relevant nonzero onent of $\mathcal{H}_0^{0\alpha}$ is

$$\mathcal{H}_0^{03} = -(1-n)\sqrt{\frac{c_1 c_2}{c_3}}\, e^{[(1-n)/2]z}$$

$$\begin{aligned}E &= \int \sqrt{-g}(T_0^0 + t_0^0) d^3x \\ &= -\frac{1}{16\pi} \frac{(1-n)^2}{2} \sqrt{\left(\frac{c_1 c_2}{c_3}\right)} \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy \int_{z_1}^{z_2} e^{[(1-n)/2]z}\, dz.\end{aligned} \tag{27}$$

also yields a nonzero value for E, the total energy of the universe.

VI. *Bianchi type IX*

The most well-known form of the Bianchi type IX model, which is the anisotropic generalization of spatially closed FRW model, is given by

$$ds^2 = dt^2 - l^2(d\psi + \cos\theta\, d\phi)^2 - m^2(\sin\psi\, d\theta - \cos\psi \sin\theta\, d\phi)^2 - n^2(\cos\psi\, d\theta - \sin\psi \sin\theta\, d\phi)^2, \quad (28)$$

where l, m, n are functions of time alone and ψ, θ, ϕ can vary from 0 to 4π, π and 2π respectively. But pseudotensorial calculations are notoriously dangerous in polar coordinates. The results are believed to be more trustworthy in cartesian coordinates. For this reason, it is worthwhile to carry out the calculations in the Bianchi IX model in cartesian coordinate system. We take the line element as

$$ds^2 = dt^2 - S^2 dx^2 - R^2 dy^2 - [R^2 \sin^2 y + S^2 \cos^2 y] dz^2 + 2S^2 \cos y\, dz\, dx, \quad (29)$$

where S and R are functions of t alone.

Calculations following the same prescription as before, the total energy comes out to be

$$E = \frac{S}{8\pi} \int \sin y\, dx\, dy\, dz \quad (30)$$

which is not equal to zero for a finite universe.

3. Discussion

Following the investigations of Rosen [5], the total energy E of the universe has been computed in the present work in some anisotropic cosmological models on the basis of Einstein pseudotensor. It has been found that although in Bianchi type I and II, E is zero, i.e. same as that obtained by Rosen [5] and also by Johri *et al* [4] but in some other examples, namely Bianchi type III, V, VI, and IX, the total energy is nonzero. So the speculation that the total energy of the universe is zero cannot in fact be proved by this type of calculations.

The present work neither intends to contest nor defend the speculation. It only shows that the pseudotensorial calculations lead to different results for different spacetime symmetries. Moreover, these objects, $\sqrt{-g}(T_j^i + t_j^i)$ do not form a tensor and thus the calculations based on it are coordinate dependent. This limitation has been noticed long back by Weyl and Pauli. For an excellent discussion on these issues, we refer to the work of Chandrashekhar and Ferrari [9]. The choice of the pseudotensor, so as to write the covariant divergence as an ordinary one, is not at all unique. Landau–Lifshitz [10] pseudotensor for instance, is another choice, which is at least symmetric unlike Einstein pseudotensor, but it has its share of problems as well. Sorkin [11] developed the idea of a conserved current in the form of Noether operator and Burnett and Wald [12] in the form of a symplectic current. These conserved currents seem to work well when the spacetime is static and thus cannot be successfully applied in a cosmological situation where the spacetime evolves with time.

If, however, the metric components depend solely on time, then it was shown by Prasanna [13] that the total energy will be zero in a finite proper volume. This result is

nsistent with our calculation as only in Bianchi type I metric all nonzero components of are functions of time alone and energy is also zero.

Very recently Aguirregabiria, Chamorro and Virbhadra [7] showed that pseudotensors Einstein, Landau–Lifshitz and many others like Tolman, Papapetrou and Weinberg e essentially the same result for the energy and energy current density components for Kerr–Schild class metric. In fact, most of these formalisms yields consistent results a Kerr–Newman spacetime, was known even earlier by the investigations of Virbhadra , 15]. As the Kerr–Schild class includes a wide variety of spacetimes, including the erior field of a radiating object, these studies are indeed very important. But this type results are not in the literature for a cosmological scenario. The coordinate independent finition of energy, given by Komar [16] and later used by many workers such as Cohen], also involves a stationary spacetime. The generalization of Komar's work by lkarni, Chellathurai and Dadhich [18] also cannot be used for a nonstatic cosmological tric. For a very brief but excellent review of different pseudotensorial formalisms, we er to the introduction of ref. [7].

To conclude, one can suggest that in order to prove (or disprove) classically, the culation that the total energy of the universe is zero, one should resort to calculations sed on tensorial objects so that these can be applied to different models with equal nfort and confidence. The method suggested by Rosen proves the conjecture ($E = 0$) y in some models and gives rise to counter examples as well.

knowledgements

e authors would like to thank N K Dadich, B R Iyer and A R Prasanna for discussions d also the anonymous referee for some helpful suggestions. One of them (SS) would e to thank the University Grants Commission, India for financial support.

ferences

1] M G Albrow, *Nature (London)* **241**, 56 (1973)
2] E P Tyron, *Nature (London)* **246**, 396 (1973)
3] F I Cooperstock, *Gen. Relativ. Gravit.* **26**, 323 (1994)
4] V B Johri, D Kalligas, G P Singh and C W F Everitt, *Gen. Relativ. Gravit.* **27**, 313 (1995)
5] N Rosen, *Gen. Relativ. Gravit.* **26**, 319 (1994)
6] A K Raychaudhuri and S Banerji, *Pramana – J. Phys.* **34**, 383 (1990)
7] J M Aguirregabiria, A Chamorro and K S Virbhadra, *Gen. Relativ. Gravit.* **28**, 1393 (1996)
8] A Einstein, *Ann. Phys.* **49**, 769 (1916)
9] S Chandrashekhar and V Ferrari, *Proc. R. Soc. London* **A435**, 645 (1991)
0] L D Landau and E M Lifshitz, *Classical theory of fields* (Pergamon, Oxford, 1962)
1] R D Sorkin, *Proc. R. Soc. London* **A435**, 635 (1991)
2] G A Burnett and R M Wald, *J. Math. Phys.* **31**, 725 (1990)
3] A R Prasanna, *Prog. Theor. Phys.* **45**, 1330 (1971)
4] K S Virbhadra, *Phys. Rev.* **D41**, 1086 (1990)
5] K S Virbhadra, *Phys. Rev.* **D42**, 2919 (1990)
6] A Komar, *Phys. Rev.* **113**, 934 (1959)
7] J M Cohen, *J. Math. Phys.* **9**, 905 (1968)
8] R Kulkarni, V Chellathurai and N Dadhich, *Class. Quant. Gravit.* **5**, 1443 (1988)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 617–622

Semi-empirical formula for Λ-binding energies in ground states of light hypernuclei

M Z RAHMAN KHAN, NASRA NEELOFER and M A SUHAIL*
Department of Physics, Aligarh Muslim University, Aligarh 202 002, India
* Department of Applied Physics, Z. H. College of Engineering and Technology, Aligarh Muslim University, Aligarh 202 002, India

MS received 5 June 1997; revised 13 October 1997

Abstract. Starting with a Λ-nucleus potential, we have obtained a semi-empirical formula, which gives a fairly satisfactory account of the ground state Λ-binding energy of light hypernuclei, if the very light nuclei are ignored.

Keywords. Semi-empirical; binding energy.

PACS No. 21.80

1. Introduction

Semi-microscopic calculations of the Λ-binding to nuclei (B_Λ) have been possible only for a few nuclei, ${}^{3}_{\Lambda}$H, ${}^{4}_{\Lambda}$H, ${}^{4}_{\Lambda}$He, ${}^{5}_{\Lambda}$He, and more approximately for ${}^{9}_{\Lambda}$Be, and ${}^{16}_{\Lambda}$O. Semi-microscopic calculations have also been carried out rather recently for a few more nuclei. On present indications, it appears unlikely, in the near future, to successfully carry out semi-microscopic calculations for many other hypernuclei. Thus, one is more or less forced to consider phenomenological calculations of B_Λ to obtain some broad insight into some of the important aspects. It is even desirable to obtain semi-empirical formulae which can give at least a rough value of B_Λ.

The Λ-nucleus potential, in the folding model with a density-dependent effective Λ-nucleon (ΛN) interaction, has been shown [1, 2] to be quite satisfactory in accounting for the ground state B_Λ data of light hypernuclei. However, all these calculations have to be carried out numerically.

By mathematical manipulations of the folded potential, several authors [3, 4] have obtained semi-empirical formulae of B_Λ for medium and heavy hypernuclei. However, none of these formulae hold for light hypernuclei for the simple reason that the core density is not of the Woods–Saxon (W–S) form as is the case for medium and heavy nuclei. No semi-empirical or even empirical formula exists for light hypernuclei because of the fact that there are different mathematical forms of the density for different light nuclei and, more importantly, simple mathematical manipulations, leading to a semi-empirical formula, do not seem to be possible for any of these forms.

So, we made a trial and error search for the Λ-nucleus potential for light hypernuclei using various analytical forms to find one that gives tolerable results by adding a

different number of protons. It was hoped that mathematical manipulation of an expression of the B_Λ may yield a semi-empirical formula. Luckily, we find that the B_Λ formula obtained using exponential form of the Λ-nucleus potential gives a semi-quantitative account of the ground state B_Λ data of light hypernuclei, ignoring only the very light hypernuclei. On general plausible grounds, we expect that our treatment would apply only to nuclei lying on the stability line and very close to it.

We obtain a semi-empirical formula for the Λ-binding energy in bound s-states of light hypernuclei using the exponential form for the Λ-nucleus potential. This formula has three adjustable parameters. After very light hypernuclei, ${}^{3}_{\Lambda}\mathrm{H}$, ${}^{4}_{\Lambda}\mathrm{H}$ and ${}^{4}_{\Lambda}\mathrm{He}$, are excluded from the fit, the formula gives a reasonable account of the B_Λ of nine light hypernuclei for which the experimental r.m.s. radius is available. With a little more relaxation on accuracy, a slightly modified formula accounts for seventeen nuclei. The formula is presented in the next section. Results and discussion is given in § 3 and conclusions in § 4.

2. A semi-empirical formula for the Λ-binding energies of light hypernuclei

We set the radial wave function $R_0(r) = \chi(r)/r$ and analytically solve the s-state radial Schrödinger equation with the exponential potential $V(r) = -V_0\,\mathrm{e}^{-r/a}$. The parameter a appearing in this potential is taken, in a plausible way, to be given as

$$a = r_0' A_c^{1/3},$$

where r_0' may be taken as a free parameter. We transform the Schrödinger equation by taking a new variable $z = \mathrm{e}^{-r/2a}$. Introducing the parameters $\alpha = +(8m_\Lambda V_0 a^2/\hbar^2)$, $\beta = +[8m_\Lambda |E| a^2/\hbar^2]^{1/2}$, where E is the energy eigenvalue, and, for definiteness, taking $\sqrt{\alpha}$ to be positive, we have

$$\chi(r) = AJ_\beta(\sqrt{\alpha}\,\mathrm{e}^{-r/2a}) + B\,Y_\beta(\sqrt{\alpha}\,\mathrm{e}^{-r/2a}). \tag{2.1}$$

When β is not an integer or zero, $Y_\beta(r)$ is simply equal to $J_{-\beta}(r)$. As $r \to \infty$, the argument of both J_β and Y_β vanishes and since Y_β is not well behaved at zero, the coefficient of Y_β has to be put equal to zero.

Further, since $R_0(r) = \chi(r)/r$, it follows that $\chi(0) = 0$. Therefore, we have

$$J_\beta(\sqrt{\alpha}) = 0.$$

For real β, the function $J_\beta(\sqrt{\alpha})$ has an infinite number of real zeroes. The nth positive zero of this function can be expressed as the following expansion [5]

$$\sqrt{\alpha} = \gamma - \frac{q-1}{8\gamma}\left[1 + \frac{Q_1}{3(4\gamma)^2} + \frac{2Q_2}{15(4\gamma)^4} + \cdots\right], \tag{2.2}$$

where $q = 4\beta^2$, $\gamma = (\beta - \frac{1}{2} + 2n)\frac{\pi}{2}$, $Q_1 = 7q - 31$, $Q_2 = 83q^2 - 982q + 3779, \ldots$.

In this form, it is a little complicated to get the energy for given potential parameters. Apart from other reasons, formula (2.2) is very unwieldy. One wants a simpler formula.

Here, we use (2.2) to derive the desired semi-empirical formula for B_Λ. In (2.2), we estimate the magnitudes of the 2nd and the 3rd terms within the square brackets, for a range of energy eigenvalues, say from $|E| = 0$ to roughly the expected value of B_Λ, with A_c

ging from 2 to 25. We take $r_0' = 0.41$ fm. This value of r_0' is, however, not crucial for our culations. Any other reasonable choice of r_0' would give very similar result. Even the ximum magnitude of the 2nd and 3rd terms, in the mass number range 2 to 25, is found be very small compared to unity and further these are of the opposite sign, making the contribution even smaller. Their contribution for nuclei in the low mass number region even less than it is for the relatively higher mass numbers. Thus, we neglect all the terms er than unity in the square bracket in equation (2.2). Then, (2.2) may be written as

$$\sqrt{\alpha} = \gamma - \frac{q-1}{8\gamma}. \tag{2.3}$$

bstituting for α, γ and β in the above equation and simplifying we get

$$B_\Lambda = \frac{\hbar^2 A_c^{-2/3}}{8m_\Lambda r_0'^2((\pi^2/2)-1)^2}\left\{(\xi_1+\xi_2) \pm \left[(\xi_1+\xi_2)^2 - (\pi^2-2) \times \left(\frac{\xi_1^2}{\pi^2} + \left(\frac{1}{2}-2n\right)\xi_2 + \frac{1}{8}\right)\right]^{1/2}\right\}^2, \tag{2.4}$$

ere $\xi_1 = \pi^2/2((1/2)-2n)$ and $\xi_2 = \pi r_0' A_c^{1/3}/2\ [8m_\Lambda V_0/\hbar^2]^{1/2}$. For the ground state, = 1. The formula (2.4) in this form does not distinguish between hypernuclei having ne A_c but different number of protons (Z). So, in order to take the Z-dependence and o the size of the core nuclei into account, we add a term $\beta' Z/r_{\rm rms}^3$, where $r_{\rm rms}$ is the .s. radius of the core nucleus being considered, to the B_Λ formula (2.4), i.e.

$$B_\Lambda = \frac{\hbar^2 A_c^{-2/3}}{8m_\Lambda r_0'^2(\pi^2/2-1)^2}\left\{(\xi_1+\xi_2) \pm \left[(\xi_1+\xi_2)^2 - (\pi^2-2) \times \left(\frac{\xi_1^2}{\pi^2} + \left(\frac{1}{2}-2n\right)\xi_2 + \frac{1}{8}\right)\right]^{1/2}\right\}^2 + \frac{\beta' Z}{r_{\rm rms}^3}. \tag{2.5}$$

e last term added above is essentially same as that taken by Rahman Khan and Shoeb who have given the justification for it. All we have done is not to take it as $\beta' Z/A_c$. r light nuclei it seems better to take $r_{\rm rms}$ directly rather than replace it with $A_c^{1/3}$. We find that formula (2.5) with the positive sign for the 2nd term within the curly acket gives reasonable values of B_Λ whereas when negative sign is chosen, quite realistic values are obtained. Thus, on this basis, we retain, in our formula (2.5) oughout, only the positive sign for the 2nd term within the curly bracket.

Results and discussion

eating V_0, r_0' and β' as adjustable parameters, χ^2-fitting is carried out for the ground te B_Λ data of the light hypernuclei with core mass number ranging from 5 to 14. cept for very light hypernuclei, ${}^3_\Lambda$H, ${}^4_\Lambda$H and ${}^4_\Lambda$He, all other light hypernuclei for which s radii of the core nuclei, in the given mass number range, are known experimentally , are included in the fit. However, predictions are made regarding the B_Λ of the cluded light hypernuclei.

Table 1. Results of χ^2-fitting of B_Λ using analytical formula (2.5). The best fit parameters for $\chi^2 = 29.04$ are: $V_0 = 103.37$ MeV, $r_0' = 0.403$ fm and $\beta' = 1.65$ MeV fm^3. B_Λ^{cal}, marked by an asterisk, are the predicted values.

Hypernuclei	A_c	$B_\Lambda^{\text{exp}} \pm \Delta B_\Lambda^{\text{exp}}$ (MeV)	B_Λ^{cal} (MeV)
$^{3}_{\Lambda}$H	2.0	0.15 ± 0.05	0.21*
$^{4}_{\Lambda}$H	3.0	2.04 ± 0.04	1.32*
$^{4}_{\Lambda}$He	3.0	2.39 ± 0.03	1.45*
$^{5}_{\Lambda}$He	4.0	3.12 ± 0.02	3.12
$^{7}_{\Lambda}$Li	6.0	5.58 ± 0.03	5.54
$^{8}_{\Lambda}$Li	7.0	6.80 ± 0.03	6.85
$^{10}_{\Lambda}$Be	9.0	9.11 ± 0.22	9.12
$^{11}_{\Lambda}$B	10.0	10.24 ± 0.05	10.23
$^{12}_{\Lambda}$B	11.0	11.37 ± 0.06	11.19
$^{13}_{\Lambda}$C	12.0	11.69 ± 0.12	12.06
$^{14}_{\Lambda}$C	13.0	12.17 ± 0.33	12.87
$^{15}_{\Lambda}$N	14.0	13.59 ± 0.15	13.59

*Predicted values.

Table 2. Results of χ^2-fitting of B_Λ from analytical formula (2.5) in which the last term is replaced by $\beta'(Z/A_c)$. The best fit parameters for $\chi^2 = 214.15$ are: $V_0 = 98.87$ MeV, $r_0' = 0.41$ fm and $\beta' = 1.00$ MeV fm^3.

Hypernuclei	A_c	$B_\Lambda^{\text{exp}} \pm \Delta B_\Lambda^{\text{exp}}$ (MeV)	B_Λ^{cal} (MeV)
$^{5}_{\Lambda}$He	4.0	3.12 ± 0.02	3.04
$^{6}_{\Lambda}$He	5.0	4.18 ± 0.10	4.35
$^{7}_{\Lambda}$Li	6.0	5.58 ± 0.03	5.77
$^{8}_{\Lambda}$He	7.0	7.16 ± 0.70	6.77
$^{8}_{\Lambda}$Li	7.0	6.80 ± 0.03	6.91
$^{8}_{\Lambda}$Be	7.0	6.84 ± 0.05	7.06
$^{9}_{\Lambda}$Li	8.0	8.50 ± 0.12	7.97
$^{9}_{\Lambda}$B	8.0	8.29 ± 0.18	8.22
$^{10}_{\Lambda}$Be	9.0	9.11 ± 0.22	9.06
$^{10}_{\Lambda}$B	9.0	8.89 ± 0.12	9.17
$^{11}_{\Lambda}$B	10.0	10.24 ± 0.05	10.05
$^{12}_{\Lambda}$B	11.0	11.37 ± 0.06	10.88
$^{12}_{\Lambda}$C	11.0	10.76 ± 0.19	10.97
$^{13}_{\Lambda}$C	12.0	11.69 ± 0.12	11.73
$^{14}_{\Lambda}$C	13.0	12.17 ± 0.33	12.44
$^{15}_{\Lambda}$N	14.0	13.59 ± 0.15	13.18
$^{16}_{\Lambda}$O	15.0	13.0	13.87

As usual the χ^2 is defined as

$$\sum_i \left[\frac{B_\Lambda^{\exp}(i) - B_\Lambda^{\text{th}}(i)}{\Delta B_\Lambda^{\exp}(i)}\right]^2,$$

where the symbols have their usual meaning, with the summation extending to all the data-points. The experimentally quoted value is taken to be $B_\Lambda^{\exp}(i) \pm \Delta B_\Lambda^{\exp}(i)$.

The calculated B_Λ corresponding to the best fit parameters, $V_0 = 103.37$ MeV, $r_0' = 0.403$ fm and $\beta' = 1.65$ MeV fm^3, are shown in table 1. Total χ^2 corresponding to the 9 data-points is 29.04. This may be taken as a fairly good agreement. Predicted B_Λ values, which are marked by an asterisk, are also shown in the table. Thus, formula (2.5) is the semi-empirical formula for B_Λ of light hypernuclei that we were looking for.

Expressing the last term of B_Λ formula (eq. (2.5)) as $\beta'(Z/A_c)$ does not give good results, presumably due to neglect of size effect which seems to have an even more important role in light hypernuclei. However, even this is qualitatively acceptable when the very light hypernuclei, ${}^{3}_{\Lambda}$H, ${}^{4}_{\Lambda}$H and ${}^{4}_{\Lambda}$He, and the well known 'troublesome' nuclei [8–10], ${}^{7}_{\Lambda}$Be and ${}^{9}_{\Lambda}$Be, are excluded from the fit. In this way they account for 17 light hypernuclei. These are given in table 2. As error bar in the experimental B_Λ data of ${}^{16}_{\Lambda}$O is not available, we take a plausible value of 5% of the experimental B_Λ as the error in the datum. Besides $\beta'(Z/A_c)$, if a constant term is also added in the formula, the χ^2 reduces significantly.

4. Conclusion

The main virtue of the exponential potential, from our point of view, is that it leads to a semi-empirical formula for the light hypernuclei, which reproduces B_Λ of many light hypernuclei fairly well. We do not offer any conjectures whether any deeper significance need be ascribed to the exponential potential. From what we have said above, the answer seems to be in the negative. Still, the calculations points to an even greater importance of size effects for light hypernuclei.

The fomula with r_{rms} (eq. (2.5)) may be used for predicting the r_{rms} of the core of the hypernuclei for which it is not experimentally known.

Acknowledgement

We are grateful to Dr. H H Ansari for discussion and providing solutions for some of the other potentials that we tried in the initial stages and failed. We are also thankful to the staff of Computer Centre, AMU, for their cooperation in using their ALPHA computer system.

References

[1] I Ahmad, M Mian and M Z Rahman Khan, *Phys. Rev.* **C31**, 1590 (1985)
[2] M Mian, *Phys. Rev.* **C35**, 1463 (1987)
[3] C Daskaloyannis, M E Grypeos, C Koutroulos and D Saloupis, *Lett. Nuovo Cimento* **42**, 257 (1985)

[4] N Neelofer, Mohammed Shoeb and M Z Rahman Khan, *Pramana – J. Phys.* **37**, 419 (1991)
[5] E Jahnke and F Emde, *Tables of functions* (Dover Publications, New York, 1945) p. 143
[6] M Z Rahman Khan and M Shoeb, *J. Phys. Soc. Jpn.* **55**, 3008 (1986)
[7] H De Vries, C W De Jager and C De Vries, *Atomic Nucl. Data Tables* **36**, 495 (1987)
[8] A R Bodmer, Q N Usmani and J Carlson, *Phys. Rev.* **C29**, 684 (1984)
[9] A R Bodmer and J W Murphy, *Nucl. Phys.* **64**, 593 (1965)
[10] A R Bodmer and Shamsher Ali, *Nucl. Phys.* **6**, 219 (1958)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 623–633

Measurement of excitation functions and mean projected recoil ranges of nuclei in ^{12}C-induced reactions on vanadium

M ISMAIL, R P SHARMA and M H RASHID
Variable Energy Cyclotron Centre, I/AF, Bidhan Nagar, Calcutta 700 064, India
Email: ismail@veccal.ernet.in

MS received 19 October 1996; revised 16 June 1997

Abstract. Excitation function and mean projected recoil ranges of nuclei produced in the ^{12}C-induced reactions on ^{51}V target were measured by conventional stacked foil and thick-target thick-recoil-catcher technique for bombarding energies $E \leq 84$ MeV for ^{12}C ion beam. The measured recoil ranges are converted to momentum transfer. Information on momentum transfer was used to get clues about some aspects of the interaction such as complete fusion which corresponds to full momentum transfer and incomplete fusion reaction mechanism. The measured excitation functions are compared with the calculation based on the statistical model which describes only equilibrium decay of the compound nucleus using the Cascade code and the geometry dependent hybrid model which describes equilibrium as well as pre-equilibrium decay of the compound nucleus using the Alice/91 code. The measured excitation functions and average ranges of the radioisotope products of the reactions ^{12}C on ^{51}V indicate that the three separate reaction mechanisms could be attributable to complete fusion of ^{12}C, incomplete fusion of ^{8}Be and incomplete fusion of ^{4}He respectively with the target. The ^{8}Be and ^{4}He are the break-up component of ^{12}C into ^{8}Be + ^{4}He. The predictions of the codes, especially the Cascade, generally agree with the measured cross-sections which could be attributed to complete fusion of ^{12}C with the target ^{51}V.

Keywords. ^{12}C-induced reactions on vanadium; stacked foil technique; isotope production; equilibrium decay; Cascade and Alice/91 codes.

PACS No. 25.60

1. Introduction

For many years now there has been a great interest to study the reaction mechanism in medium energy heavy-ion induced reactions. Most of the earlier studies have concentrated on the energy, angular momentum and charge distribution of the products emitted in heavy-ion reactions, while others are interested in multifragmentation which are intimately connected with that of complex or intermediate mass fragments. In recent years there has been considerable interest to study fusion and incomplete fusion (ICF) in heavy ion reactions at projectile energies in the range of 5–10 MeV/amu [1]. The motivation for this work was to use the ^{12}C ions and other ion beams such as ^{7}Li, ^{14}N and ^{16}O from BARC–TIFR Pelletron Accelerator Facility to study the fusion and incomplete fusion (ICF) reaction mechanism by measuring the cross-section and recoil ranges of

radioactive products which could be converted to linear momentum transfer to the compound nucleus.

In the case of fusion reaction the highly excited nuclear system decays by evaporating low energy nucleons and alpha particles at the equilibrium stage. While in the case of ICF reaction only a part of the projectile fuses with the target nucleus and the other part moves in beam direction with almost the same velocity as that of incident ion beam. The excitation function and mean projected recoil ranges of nuclei produced in the ^{12}C-induced reactions on ^{51}V target were measured by conventional stacked foil and thick target thick recoil catcher technique for bombarding energies $E \leq 84$ MeV ^{12}C ion beam. The work reported here also relates to the measurements of the recoil ranges of nuclei produced, to get experimental information about the momentum transfer associated with specific final nuclei. The linear momentum transfer characteristics of reactions supplement the energy partitioning characteristics of the reactions. The average linear momentum transfer characterizes the global features of the projectile target interaction and provides a convenient indication of the reaction mechanism. It will be largest for complete fusion reactions and reduced for direct (incomplete fusion), pre-equilibrium and break-up processes in which much of the incident momentum is carried away by the fast particles in the early stages of the reaction. Experiments have been performed at the BARC–TIFR (14-UD Pelletron) Medium Energy Heavy Ion Accelerator Facility at Bombay. Several techniques have been devised to measure the momentum of the recoil nuclei. Measurement of recoil velocity of the detected heavy products can be used to identify the fusion modes i.e. the complete fusion or incomplete fusion (ICF). Attempts to measure recoil velocity have been made in time of flight studies; however only the mass of the residue is identified and not its atomic number. If the heavy residue is a radio isotope, the original recoil velocities could be inferred from its range measurement in a stopping medium. This is a long established way of investigating heavy-ion reactions [2–4].

In the present work the linear momentum transfer characteristic in the ^{12}C-induced reactions on ^{51}V, was obtained employing the classical method of thick-target thick-catcher recoil range measurements [2–4]. Several excitation functions for the ^{12}C-induced reactions on ^{51}V measured by using the stacked foil technique are presented. The excitation functions of the radioactive products observed in reactions contain some information about the mechanism of the interaction of the ^{12}C-induced reactions on ^{51}V and we will be able to decipher some features of the partitioning of the incident energy into that carried away by outgoing fast particles at the pre-equilibrium stage (in contrast to evaporation), and that left as nuclear excitation by using the code Alice/91 [5]. In this work calculations in the framework of the equilibrium statistical model using the codes Cascade [6] and pre-equilibrium model using the code Alice/91 [5] were performed and the results are compared with experimental excitation functions. Various reaction mechanisms, supported by recoil range measurements, are indicated as contributing to the production of the radioactive nuclei.

2. Experimental procedure

Excitation functions and mean projected recoil ranges of the nuclei produced in the ^{12}C-induced reactions on ^{51}V were measured using the absolute yields of characteristic γ-rays

pertaining to the decay of each radioactive residual nuclei by off-line γ-ray measurements. The targets were obtained by rolling vanadium foils using the rolling machine at TIFR, Bombay. The thickness of the targets for excitation function measurements was about $\cong 2.0\,\mu m$ whereas for the range measurements the thickness was about $\cong 8.0\,\mu m$. The stacks were formed by placing alternatively 4.34 μm aluminium foils as degraders of ^{12}C energy over the targets. Then the stacks were irradiated in a chamber specially constructed for this purpose by BARC, Bombay and having the facility to suppress electron in the Faraday cup by applying negative bias. The beam spot on the targets was limited to 5.0 mm diameter by using a stainless steel collimator in front of the targets. There were 13 targets in the stack for the excitation function measurement and 3 to 4 targets in the stack for recoil range measurements. The stacks were exposed to the analysed beam from the 14UD tandem pelletron accelerator facility at TIFR, Bombay. The beam current on the targets was kept below 200 enA. The total ^{12}C-beam was collected and measured using a calibrated current integrater.

The average thickness of the target and degrader foils was determined by weighing. Each foil was cut out into a square shape and pasted on an annular aluminium holder having 21.0 mm as the outer diameter and 15 mm as the inner diameter.

The efficiency calibrations of the detector were done with a standard ^{152}Eu radioactive source. The efficiency of the detector was interpolated to the required energy value from the measured efficiency curve. However to improve the interpolation, the efficiency curve was fitted to a function of the form

$$\varepsilon(E_\gamma) = \sum_{i=1,4} A_i * (E_\gamma)^{i-1} + B_1 \exp(-B_2 * E_\gamma) \tag{1}$$

by a non-linear least square method (for details see ref. [7]) where A_i and B_i are constants to be determined by least-square fit and E_γ is the γ-ray energy. The fit improves the accuracy of interpolation considerably.

The mean beam energy at half thickness in each foil of a stacked foil assembly was calculated from energy degradation of the initial beam energy using the coefficients obtained from fitting the stopping power data for different materials. The α-stopping power tables of Williamson *et al* [8] were used for fitting. Then the stopping power S for a given combination of stopping medium and heavy-ion beam and velocity is calculated by means of the well-known scaling law

$$S/(\gamma Z_1)^2 = S_{\text{ref}}/Z_{\text{ref}}^2 \tag{2}$$

in which (γZ_1) is the heavy-ion effective charge (Z_1 being its atomic number) and S_{ref} is the stopping power of the same medium for a reference ion (α-ion) of the same velocity and of effective charge Z_{ref}. The tabulation is then generated based on the arguments of the effective charge parameter $\gamma\,(Z_1, E/A, Z_2)$ given in [9]. The stopping power generated in this way is very similar to that generated by the code TRIM.

The production cross-sections were determined using the absolute yields of characteristic γ-rays belonging to each final nucleus. The γ-rays emitted by the irradiated targets and their associated catcher foils were detected with a HPGe detector available at TIFR, Mumbai as well as VECC, Calcutta. In most of the cases the γ-rays used in the yield determination (listed in table 1) stood out very prominently in the spectra and did

Table 1. Half-lives, γ-energies and branching ratios of the γ-rays and Q-values for C-induced reactions on V.

Nuclide	Half-lives	E_γ (keV)	I_γ (abs)	Reaction	Q-values (keV)
^{24}Na	14.96 h	1368.6	1.0	^{27}Al(C, $2\alpha 4p3n$)^{24}Na	−66997.4
				^{27}Al(C, $3\alpha 2pn$)^{24}Na	−38701.5
				^{27}Al(C, C $2pn$)^{24}Na	−31426.7
				^{51}V(C,^{37}Clpn)^{24}Na	−27378.9
^{58}Co	70.76 d	810.80	0.994	^{51}V(C, $2p3n$)^{58}Co	−31147.6
				^{51}V(C, αn)^{58}Co	−2851.7
^{57}Co	271.7 d	122.10	0.856	^{51}V(C, $2p4n$)^{57}Co	−39720.7
				^{51}V(C, $\alpha 2n$)^{57}Co	−11424.9
^{56}Co	78.76 d	846.80	0.999	^{51}V(C, $2p5n$)^{56}Co	−51096.6
				^{51}V(C, $\alpha 3n$)^{56}Co	−22800.7
^{54}Mn	312.2 d	835.00	1.000	^{51}V(C, $4p5n$)^{54}Mn	−66158.7
				^{51}V(C, $\alpha p3n$)^{54}Mn	−37862.8
				^{51}V(C, $2\alpha n$)^{54}Mn	−9566.9

not pose any identification problem. The γ-ray spectra from the HPGe spectrometers were recorded on 3.5″ diskettes by using PC-based data acquisition system. The spectra were later analysed by personal and SUPER-32 computers. The methods of analysis of the γ-ray spectra and the efficiency calibrations of the detector were the same as reported in ref. [10, 7].

The nuclear data necessary for the evaluation of the cross sections are presented in table 1. The half-lives of the radioactive atoms are taken from the chart of nuclides, the γ-rays energies and branching ratios are taken from the table of isotopes [12]. In table 1 only those γ-rays are listed which were chosen for the calculation of the cross-sections. Q-values were calculated using the atomic mass table of Wapstra and Audi [13].

2.1 *Recoil range measurements*

Several techniques have been devised to measure the momentum of the recoiling nuclei. In the present work the linear momentum transfer characteristics in the ^{12}C-induced reactions on ^{51}V, were obtained employing the classical method of thick-target thick-catcher recoil range measurement technique ([2], [3] and [4]). If the target and catcher foils are perpendicular to the beam axis and the target thickness is larger than the maximum observable recoil range and let x be the fraction recoiling forward, then $(1-x)$ is the corresponding fraction recoiling backwards, N the number of residual nuclei produced per unit length of the target thickness, R_F and R_B the forward and backward ranges, T the target thickness and A_F, A_B and A_T the activities induced in the forward and backward catchers and in the target. Then it easily follows that

$$A_F = xNR_F; \quad A_B = (1-x)NR_B; \quad A_T = NT - A_F - A_B. \tag{3a–c}$$

ence

$$A_F/(A_F + A_T) = xR_F/[T - R_B(1-x)]. \tag{3d}$$

ince x is close to unity, one can neglect the last term in the denominator $(1-x) \cong 0$ hich is small in comparison with the target thickness and then one obtains

$$R_F = T \cdot A_F/(A_F + A_T). \tag{4}$$

his relation has been used in this work to measure the mean forward range in the ^{12}C-duced reactions on ^{51}V.

In order to obtain information about the linear momentum of the recoil products, their nges were converted to velocities parallel to the beam direction (V_p) with the help of nge-energy relation $d\varepsilon/d\rho$ given by Lindhard, Scharff and Schiott theory [14]. The opping power $d\varepsilon/d\rho$ is the sum of two terms

$$d\varepsilon/d\rho = (d\varepsilon/d\rho)_e + (d\varepsilon/d\rho)_n. \tag{5}$$

or the electronic energy loss, the expression $(d\varepsilon/d\rho)_e = k\sqrt{\varepsilon}$ given by Lindhard and charff [15] is used. For the energy loss due to nuclear collision the expression $d\varepsilon/d\rho)_n = -c\log\{\varepsilon^\alpha/a + b/\varepsilon^\beta\}$ with $c = 0.2$, $\alpha = 1.215$, $a = 70.0$, $b = 0.002$ and

Table 2. Experimental and calculated ranges for the products in ^{51}V + ^{12}C reaction.

$\pm \Delta E$ MeV)	Recoil energy (MeV)	Range (μm) and recoil energy of the products				
		^{58}Co	^{57}Co	^{56}Co	^{56}Mn	^{54}Mn
0.8 ± 3.2	15.4	3.3 (3.8 ± 0.3) 14.2*	3.3 (2.5 ± 0.2) 13.9*	3.2 (2.6 ± 0.2) 13.7*	3.4 (2.7 ± 0.2) 13.7*	3.3 (1.7 ± 0.1) 13.2*
2.1 ± 3.5	13.7	3.1 (3.0 ± 0.2) 12.6*	3.0 (3.0 ± 0.2) 12.4*	2.93 (2.5 ± 0.2) 12.2*	3.2 (2.3 ± 0.2) 12.2*	3.0 (2.1 ± 0.2) 11.8*
6.3 ± 3.7	12.6	2.9 (3.4 ± 0.3) 11.6*	2.8 (3.1 ± 0.3) 11.4*	2.8 (2.3 ± 0.2) 11.2*	3.0 (2.1 ± 0.2) 11.2*	2.9 (2.3 ± 0.2) 10.8
2.7 ± 3.8	11.9	2.7 (2.1 ± 0.2) 11.0*	2.7 (2.7 ± 0.2) 10.8*	2.7 (2.8 ± 0.2) 10.6*	2.8 (1.4 ± 0.1) 10.6*	2.7 (1.8 ± 0.1) 10.2*
6.2 + 4.0	10.7	2.5 (3.3 ± 0.3) 9.9*	2.5 (2.6 ± 0.2) 9.7*	2.4 (1.4 ± 0.1) 9.5*	2.6 (0.8 ± 0.1) 9.5*	2.5 (1.7 ± 0.1) 9.2*
4.9 ± 4.7	8.6	2.1 (4.0 ± 0.3) 7.9*	2.1 (1.8 ± 0.1) 7.7*	2.0 7.6*	2.2 7.6*	2.1 (0.7 ± 0.1) 7.3*

he measured recoil ranges are within ( ) and the corresponding recoil energies are marked with *). The calculated ranges are having ≅ 15% error due to errors in the stopping power parameter. ee eq. (5).

$p = 0.815$. In the recoil range limits which are relevant to the present work, the ranges are almost exactly proportional to the recoil velocity as shown above as well as in Northcliff and Schilling [16, 17]. It was shown in [18] that under these conditions, and with the assumption of an isotropic distribution of the evaporation vector, the evaporation contribution to the thick target recoil ranges may be neglected to the first order. The correction to (V_p) arising from the angular distribution of the recoils $\langle \cos \phi \rangle$, due to atomic collisions, was calculated by the expression derived by Blaugrund [19], where ϕ is the average scattering angle due to nuclear collision. The constants used in the expressions for $(\mathrm{d}\varepsilon/\mathrm{d}\rho)_e$ and $(\mathrm{d}\varepsilon/\mathrm{d}\rho)_n$ are having $\cong 15\%$ errors, thereby introducing $\cong 15\%$ errors in the calculated range values also.

The technique for the range measurement mentioned above assumes uniform production cross-section for the nuclei throughout the target width in energy units so that the activities A_F, A_B and A_T are exactly proportional to forward range, backward range and total target thickness. In the present work range measurements for the ^{12}C-induced reactions on ^{51}V target are shown in table 2. The thickness of ^{51}V targets correspond to $\cong 6.4$ MeV in energy units for 84 MeV ^{12}C-ion and $\cong 09.5$ MeV for the 44 MeV ^{12}C-ion. Therefore, in the present case (^{51}V target) the cross-section at the two edges differ appreciably compared to the average cross-section over the target thickness. If the bombarding energy is at the rising part of the excitation function as seen from the high energy side then the range extracted by this method will be larger than the actual range whereas if the bombarding energy is at the declining part of the excitation function then the opposite will be true i.e. the extracted range will be smaller than the actual range. This is the reason why some of the extracted ranges are higher and others smaller than the calculated ranges. However the extracted ranges help in some cases to distinguish between compound nucleus by fusion reaction and incomplete fusion reaction including direct reaction mechanisms.

2.2 *Cross-section determination*

The number of observed decays Z per unit time is related to the total number of decays Z_0 per unit time by

$$Z_0 = Z/[\varepsilon(E_\gamma) * I_\gamma(\mathrm{abs})], \tag{6}$$

where $\varepsilon(E_\gamma)$ is the detector efficiency and $I_\gamma(\mathrm{abs})$ is the absolute γ-ray abundance per decay. The total number of decays, Z_0, is related to total reaction yield R for simple decays (simple decays correspond to direct production of radio isotopes by the nuclear reactions and we have used only such decays in all the measurements reported here) by the relation

$$R = Z_0 \cdot \exp(\lambda T_2)/[(1 - \exp(-\lambda T)) \cdot (1 - \exp(-\lambda T_1))/\lambda T_1], \tag{7}$$

where λ is the decay constant, T_1 the irradiation time and T_2 the cooling time after the end of irradiation and the beginning of the measurement and T is the counting time. R is related to the cross-section σ by the relation

$$\sigma = R/[(N_A \cdot \delta x) \cdot \mathcal{I}], \tag{8}$$

Table 3. Experimental cross-section for ^{12}C-induced reaction with ^{51}V.

Energy $\pm\Delta E$ (MeV)	Cross-sections for the products (mb)				
	Co-57	Co-58	Co-56	Mn-54	Na-24
83.2	48.4	122.0	23.3	79.3	13.3
±0.8	±3.9	±9.8	±1.9	±6.3	±1.1
79.7	64.8	185.9	44.4	109.1	11.7
±0.8	±5.2	±14.9	±3.6	±8.7	±0.9
76.0	68.4	170.4	47.8	119.6	10.5
±0.9	±5.5	±13.6	±3.8	±9.6	±0.8
72.3	73.1	138.4	51.3	131.8	8.0
±0.9	±5.8	±11.1	±4.1	±10.5	±0.6
68.5	102.2	114.1	60.7	155.1	6.0
±0.9	±8.2	±9.1	±4.9	±12.4	±0.5
64.4	132.7	72.9	52.9	163.7	4.5
±1.0	±10.6	±5.8	±4.2	±13.1	±0.4
60.2	163.0	48.6	37.7	154.4	2.9
±1.0	±13.0	±3.9	±3.0	±12.4	±0.2
55.7	200.7	53.5	22.5	135.9	2.3
±1.1	±16.1	±4.3	±1.8	±10.9	±0.2
51.1	180.9	67.3	7.7	79.3	1.8
±1.1	±14.5	±5.4	±0.6	±6.3	±0.1
46.0	166.8	93.4	0.0	36.9	2.5
±1.2	±13.3	±7.5		±3.0	±0.2
40.5	94.9	90.0	0.0	12.3	1.7
±1.3	±7.6	±7.2		±1.0	±0.1
34.4	25.4	52.2	0.0	0.0	3.1
±1.5	±2.0	±4.2			±0.2

ere N_A is the number of atoms/cm^3 of the target material, δx is the thickness of the foil cm) and $\mathcal{I}$ is the total number of incident particles during irradiation (calculated from l charge measured by current integrator).

Experimental results

able 3 and figures 1a to 1d our experimental results for the production of ^{54}Mn, ^{56}Co, o, ^{58}Co and ^{24}Na radio-nuclides via ^{12}C-induced reactions on ^{51}V are summarized. radio-isotope ^{24}Na could also be produced by ^{12}C induced reaction on ^{27}Al which is d as catcher foils in the stacks. The possible Q-values are listed in table 1. The ertainty given for the energy values includes those of target thickness and beam rgy resolution (±5.0 keV) only. The solid lines and the dashed lines are respectively Cascade code [6] and Alice/91 [5] fits. In table 3 the experimental cross-sections for reaction $^{12}C + {}^{51}V$ are presented respectively along with absolute errors of $\cong 8\%$. The olute error consists of uncertainties due to target foil thickness (±2%), the beam rent integration (±2%), the detector efficiency (±5%) and the analysis of the γ-ray ctra (statistical uncertainty) generally ($\leq 2\%$). The uncertainties caused by the large e of the irradiation area and the non-uniformities of the target contribute about 5% to

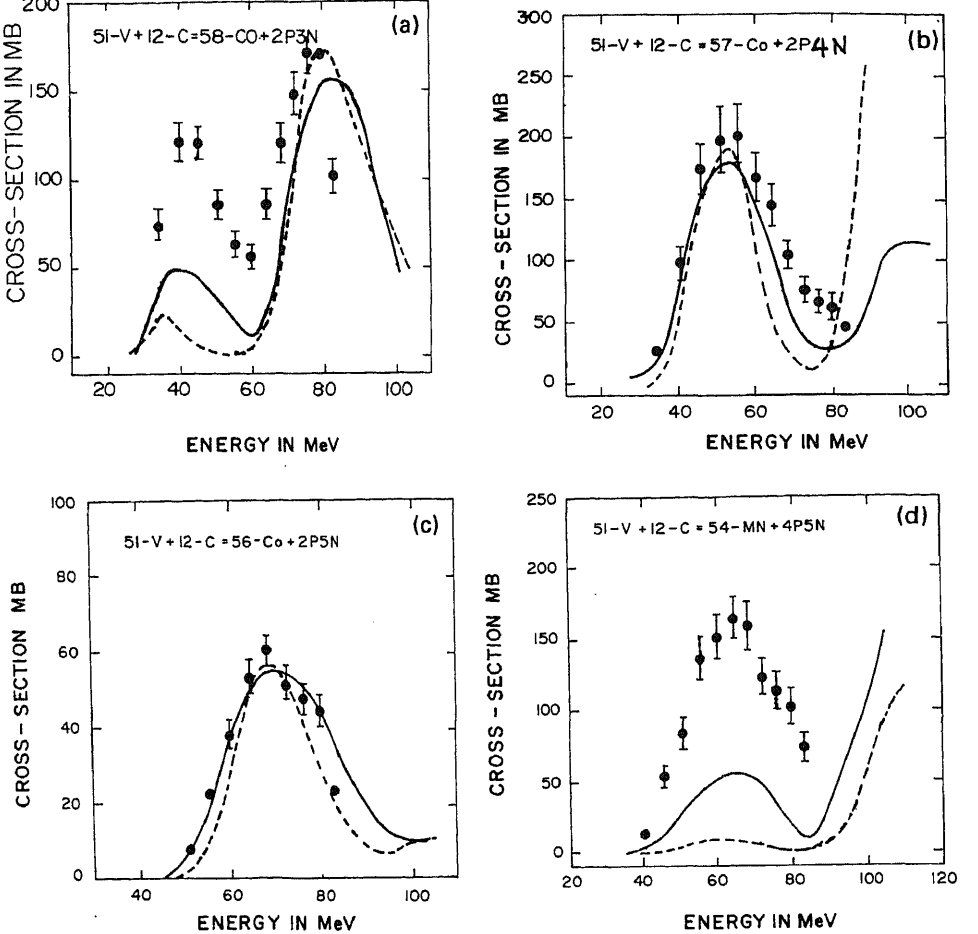


Figure 1. The total residual production cross-section in mb for the reaction (**a**) ^{51}V (^{12}C, $2p3n$) ^{58}Co (•), (**b**) ^{51}V (^{12}C, $2p4n$) ^{57}Co (•), (**c**) ^{51}V (^{12}C, $2p5n$) ^{56}Co (•) and (**d**) ^{51}V (^{12}C, $4p5n$) ^{54}Mn (•) are plotted as a function of ^{12}C-particle bombarding energy. Solid lines are the Cascade model calculation and dashed lines are Alice/91 calculation with $N_0 = 12$, $N_n = 6$, $N_p = 6$ and $N_h = 0$, normalized to experimental values at the maximum cross-section point with normalization constant $N = 1.00$ for all the solid line curve and for the dashed line curves (**a**) $N = 0.5$, (**b**) $N = 1.0$, (**c**) $N = 0.3$ and (**d**) $N = 1.0$.

the average error of the cross-section. However, the above mentioned average error values do not include the uncertainties of the nuclear data used in the analysis.

4. Comparison with theoretical predictions

A rigorous theory of nuclear reaction for which calculations can be performed to describe mass, angular and energy distribution data from intermediate energy collision with

nplex particle is presently not available. A simpler compound statistical model is ierally applicable below 7 MeV/nucleon owing to very low probability for pre-uilibrium and direct reaction processes [20]. However, the energy range of present rk is also ≤ 7 MeV/nucleon only. Hence our discussion will be based mostly on npound statistical models. It is well-known that the compound statistical model gives a rect overall description of the excitation functions and particle energy spectra in clear reactions at medium energies. However, the calculations fail to account for all ails such as the exact position of the maximum or the slope of the ascending and scending part of the excitation functions. We have investigated the nuclear mechanism the nuclear reaction $^{12}C + ^{12}V$ by using the two computer codes Cascade [6] and ce/91 [5]. The Cascade code is based on the assumption that the projectile and target m a compound nucleus in statistical equilibrium and Hauser–Feshbach formula ether with the statistical nuclear model are applied in order to calculate the intensities the various decay chains and thus the excitation functions of the reactions. The ce/91 code describes the process of equilibrium evaporation of particles and γ-rays in ms of the Weisskopf and Ewing model [21] and pre-equilibrium reaction mechanism ording to the hybrid and geometry dependent hybrid model. The high energy part the excitation functions are dominated by pre-equilibrium reaction mechanism ereas the low energy parts are dominated by evaporation with its characteristic peak as n in [7, 10]. Besides evaporation of neutrons and protons, clusters such as deuteron and particles can be considered. The nuclear masses were taken from Wapstra and Audi] mass table if available, otherwise calculated from the Mayers and Swiatecki mass mula [22], liquid drop masses with pairing. The inverse cross-sections were calculated ng the optical model subroutine of Alice/91 [5], where the optical model parameters re those of Becchetti and Greenlees [23]. In the present work our excitation functions calculated on the basis of (i) statistical model which describes only equilibrium decay ng the code Cascade [6] and (ii) geometry dependent hybrid model [24] using the gram Alice/91 [5] on the super-32 computer at our centre. In figures 1a to 1d, the ed circles are measured cross-sections. The theoretical values shown in the figures 1a 1d are multiplied by a factor so as to match the experimental data at the maximum ss-section point. The multiplying factors are given in the figure captions. The ltiplying factors also indicate the quality of fit between experimental and theoretical ues. The dashed curves and solid curves are the calculated cross-section based on the les Alice/91 and Cascade respectively. Figures 1a to 1c show that the excitation ctions for the $2pxn$ or αxn channels are predicted reasonably well by Cascade code ereas Alice/91 code fits only ^{56}Co excitation function fairly well. However there is exception i.e. the first peak in the excitation function for the production of ^{58}Co ich the codes underpredict. This is mostly due to αn channel as seen in the excitation ction for the reaction $^{59}\mathrm{Co}(\alpha, \alpha n)^{58}\mathrm{Co}$ (see ref. [11]). Figure 1d shows the excitation ction for the $2\alpha n$ or $4p5n$ channels for the production of ^{54}Mn. Both the codes derpredict the production cross-sections. These possibly indicate that these radio-topes ^{58}Co and ^{54}Mn are not only produced by α-evaporation but also by another chanism possibly by break-up of ^{12}C into α and ^{8}Be. Then the fusion of ^{8}Be and poration of a neutron to produce ^{58}Co and the fusion of α and evaporation of a neutron produce ^{54}Mn radio isotopes. The measured average ranges also support the omplete fusion reaction mechanism. In Cascade and Alice/91 codes evaporations of

only n, p and α particles are being taken into account. That is why they underpredict the incomplete fusion of ^{8}Be and ^{4}He with the target ^{51}V. However, the predictions of these codes, especially the Cascade, generally agree with the measured cross-sections which could be attributed to complete fusion of ^{12}C with the target ^{51}V.

5. Conclusion

A consistent set of 4 excitation functions has been measured for ^{12}C-induced reactions on ^{51}V target. The measured excitation functions and average ranges of the radioisotope products of the reactions ^{12}C on ^{51}V give some indication that the three separate reaction mechanisms which could be attributed to complete fusion of ^{12}C, incomplete fusion of ^{8}Be and incomplete fusion of ^{4}He respectively with the target. The ^{8}Be and ^{4}He are the break-up component of ^{12}C into ^{8}Be+^{4}He. It can be seen that the predictions of the codes, especially the Cascade, generally agree with the measured cross-sections which could be attributed to complete fusion of ^{12}C with the target ^{51}V.

Acknowledgements

The authors thank the staff of Pelletron at TIFR for help. They also thank Mr P P Burte for technical help and Dr M Blann for supplying the Alice/91 code.

References

[1] B S Tomar, A Goswami, A V R Reddy, S K Das, P P Burte, S B Manohar and Satyaprakash, *Z. Phys.* **A343**, 223 (1992)
B S Tomar, A Goswami, A V R Reddy, S K Das, P P Burte and S B Manohar, *Phys. Rev.* **C49**, 941 (1994)
[2] N Sugarman, M Campos and K Wielgoz, *Phys. Rev.* **101**, 388 (1956)
[3] J M Alexander, in *Nuclear Chemistry* edited by L Yaffe (Academic, New York, 1968) vol. 1, p. 273
[4] J M Alexander and G N Simonoff, *Phys. Rev.* **B133**, 93 (1964)
[5] M Blann, Alice/1991 – *A Evaporation Code* (Lawrence Livermore National Lab, University of California, Livermore UCID-20169, 1985)
[6] F Puhlhofer, *Nucl. Phys.* **A280**, 267 (1977)
[7] M Ismail, *Phys. Rev.* **C41**, 87 (1990)
[8] C F Williamson, J P Boujot and J Picard, CEA-R 3042 (1966)
[9] F Hubert, R Bimbot and H Gavin, *Atomic Data Nucl. Data Tables* **46**, 1 (1990)
[10] M Ismail and A S Divatia, *Pramana – J. Phys.* **30**, 193 (1988)
M Ismail, *Pramana – J. Phys.* **32**, 605 (1989)
[11] M Ismail, *Pramana – J. Phys.* **40**, 227 (1993)
R Michel and G Brinkmann, *Nucl. Phys.* **A338**, 167 (1980)
[12] C M Lederer and V S Shirley, *Table of Isotopes*, 7th edition (New York, John Wiley, 1978)
[13] A H Wapstra and G Audi, *Nucl. Phys.* **A432**, 1 (1985)
[14] J Lindhard, M Scharff and H E Schiott, *Matt. Fys. Medd. Dan. Vid. Selsk.* **33**, 14 (1963)
[15] J Lindhard and M Scharff, *Phys. Rev.* **124**, 128 (1961)
[16] L C Northcliffe, *Annu. Rev. Nucl. Sci.* **13**, 67 (1963)
[17] L C Northcliffe and R F Schilling, *Nucl. Data Tables* **A7**, 233 (1970)
[18] L Winsberg and J M Alexander, *Phys. Rev.* **121**, 518 (1961)
[19] A E Blaugrund, *Nucl. Phys.* **88**, 501 (1966)

[0] R Birkland, L E Tubbs, J R Huizenga, J N De and D Sperber, *Phys. Rep.* **56**, 108 (1979)
[1] V F Weisskopf and D H Ewing, *Phys. Rev.* **57**, 472 (1940)
[2] W D Myers and W J Swiatecki, *Nucl. Phys.* **81**, 1 (1966)
W D Myers and W J Swiatecki, *Ark. Fys.* **36**, 343 (1967)
[3] F D Becchetti and F D Greenlees, *Phys. Rev.* **182**, 1190 (1969)
[4] M Blann, *Phys. Rev. Lett.* **27**, 337 (1971)
M Blann, *Phys. Rev. Lett.* **28**, 757 (1972)

herical aberration and its correction using Lie algebraic thods

VINDAN RANGARAJAN* and MINITA SACHIDANAND
rtment of Mathematics, Indian Institute of Science, Bangalore 560 012, India
o at Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012, India
l: rangaraj@math.iisc.ernet.in

eceived 7 August 1997; revised 18 September 1997

ract. We discuss three methods to correct spherical aberration for a point to point imaging m. First, results obtained using Fermat's principle and the ray tracing method are described y. Next, we obtain solutions using Lie algebraic techniques. Even though one cannot always n analytical results using this method, it is often more powerful than the first method. The t obtained with this approach is compared and found to agree with the exact result of the first od.

vords. Spherical aberration; Lie algebraic methods.

S Nos 42.15; 02.20

ntroduction

is paper we discuss three different computational methods for correction of spherical ration in an optical imaging system. The imaging system considered is a lens that sses rays originating from a single initial point onto a final point. The final point is at ame distance from the centre of the lens as the initial point.

the above system is assumed to be ideal, as in figure 1, it would take all light rays nating from point a and focus them onto the focal point b irrespective of the angles e rays make with the optic axis. If the system is nonideal (as it would be in general), above situation is obtained only in the Gaussian or paraxial approximation. In the Gaussian regime, rays are displaced away from the focal point. In particular, if rical aberration [1] is present, a given ray is displaced in the transverse direction y from the paraxial focus. The magnitude of this displacement is dependent on the e the ray makes with the optic axis.

ur system is taken to be symmetric about a plane through the origin perpendicular to optic axis. This enables us to split the system in two right down the middle, and to ider the equivalent but simpler system shown in figure 2. Our goal is to find the tion for the curved surface of the lens (in figure 2) such that all rays come to a sharp s at the point b irrespective of the angles they make with the optic axis. Then, this ging system would be free of on-axis spherical aberration.

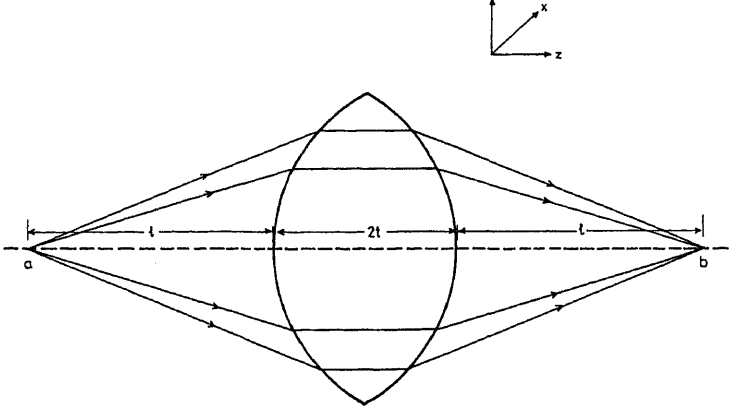


Figure 1. An ideal point-to-point imaging system. It takes all light rays originating at a point a and focusses them onto the focal point b. Points a and b are equidistant (at a distance $t + l$) from the centre of the lens. The thickness of the lens along the optic axis is $2t$. Our choice of coordinate system is also indicated in the figure.

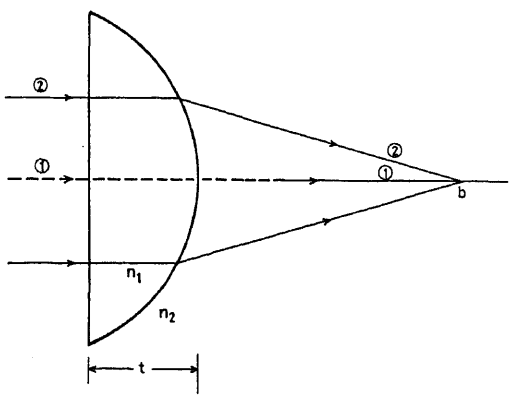


Figure 2. A simpler imaging system equivalent to the original one given in figure 1. Refractive index $n_2 = 1$ in our case.

In the next section, we directly use Fermat's principle to derive the known analytic expression [1] for the surface of the lens. However, such analytic expressions are possible only for special systems. In the subsequent sections, we compute the surface using more powerful methods that can be used to obtain solutions systematically for general optical systems. In the present case, the exact solution allows us to benchmark solutions obtained using the latter methods. We also briefly describe the result using the method of ray tracing.

In § 3, we introduce the Lie algebraic treatment of optical systems. This method is applied to the problem at hand in § 4, and the results are shown to agree with the exact solution. We give our conclusions in the final section.

2. Exact solution using Fermat's principle

In this section, we derive the equation for the surface of the lens using Fermat's principle. For our optical system (see figure 2), this principle states that rays 1 and 2 should have the same optical path length. Thus (in the $y-z$ plane)

$$A = n(t+z) + \sqrt{(l-z)^2 + y^2} = nt + l, \tag{2.1}$$

where n is the index of refraction of the lens, t is its thickness along the optic axis and l is its focal length.

Solving the above equation, we obtain z as a function of y (we pick the solution where z decreases as y increases, this being the physically appropriate solution for our problem):

$$z = \frac{l}{n+1}\left[1 - \sqrt{1 + \frac{(n+1)y^2}{(n-1)l^2}}\right]. \tag{2.2}$$

In fact, this is an equation for a hyperbola. Expanding the square root, we get

$$z = -\frac{y^2}{2l(n-1)} + \frac{(n+1)y^4}{8l^3(n-1)^2} - \frac{(n+1)^2y^6}{16l^5(n-1)^3} + \frac{5(n+1)^3y^8}{128l^7(n-1)^4} + \cdots. \tag{2.3}$$

Thus we have succeeded in obtaining the exact shape of a lens such that our point-to-point imaging systems are free of on-axis spherical aberration. This solution will enable us to verify the correctness of results obtained using the Lie algebraic methods.

One can rederive the equation for the surface of the lens using the method of ray tracing. Employing Snell's law, the following differential equation satisfied by the surface of the lens can be obtained as

$$\frac{dz}{dy} = \frac{y}{(l-z) - n[(l-z)^2 + y^2]^{1/2}}. \tag{2.4}$$

The above differential equation can be solved numerically to obtain the equation for the surface of the lens. It can also be easily verified that the exact solution derived earlier [cf. eq. (2.2)] satisfies this differential equation.

3. Lie algebraic techniques in optics

In this section, we consider the application of Lie algebraic techniques to a general optical system [2, 3] with the optic axis oriented along the z-direction. We assume the optical system to be axially symmetric about the z-axis and also symmetric with respect to reflections through some plane containing the z-axis.

The optical path length A of a general ray going from the plane z^i to the plane z^f is defined as follows:

$$A = \int_a^b n(r)\mathrm{d}s, \tag{3.1}$$

where $n(r)$ is the index of refraction at a point r and s is the distance parameter along a curve joining points a and b. Taking z to be the independent variable, the distance

function ds (in A) now becomes

$$ds = (dx^2 + dy^2 + dz^2)^{1/2} = (1 + x'^2 + y'^2)^{1/2} dz, \tag{3.2}$$

where the prime denotes differentiation with respect to z. Thus

$$A = \int_{z^i}^{z^f} n(x, y, z)(1 + x'^2 + y'^2)^{1/2} dz. \tag{3.3}$$

Comparing this with the action functional in mechanics [4], the integrand in the expression for A can be identified as the Lagrangian L for our system

$$L = n(x, y, z)(1 + x'^2 + y'^2)^{1/2}. \tag{3.4}$$

We can also define the conjugate momenta:

$$\begin{aligned} p_x &= \frac{\partial L}{\partial x'} = n(x, y, z) \frac{x'}{[1 + x'^2 + y'^2]^{1/2}}, \\ p_y &= \frac{\partial L}{\partial y'} = n(x, y, z) \frac{y'}{[1 + x'^2 + y'^2]^{1/2}}. \end{aligned} \tag{3.5}$$

Apart from the factor of n, these are just the direction cosines of the ray. Using these, we obtain the Hamiltonian in the usual way:

$$H = -[n^2(x, y, z) - p_x^2 - p_y^2]^{1/2}. \tag{3.6}$$

In Lie perturbation theory, instead of solving the Hamilton's equations of motion directly, one solves the problem in the following map form

$$w^f = \mathcal{M} w^i, \tag{3.7}$$

where $\mathcal{M}$ is a nonlinear symplectic map mapping the initial values of the phase-space 4-vector $w = (x, y, p_x, p_y)$ to its final values. For further details see ref. [5]. There is a standard procedure to obtain $\mathcal{M}$ from the Hamiltonian which is described elsewhere [6]. It can be further shown that $\mathcal{M}$ can be factorized as follows:

$$\mathcal{M} = e^{:f_2:} e^{:f_3:} e^{:f_4:} \cdots e^{:f_n:} \cdots, \tag{3.8}$$

where f_m's are homogeneous polynomials of degree m in the phase space variables. The quantity $:f_m:$ is the Lie operator associated with the function f_m and is defined by its action on a general phase-space function g as follows:

$$:f_m:g = [f_m, g], \tag{3.9}$$

where $[,]$ denotes the usual Poisson bracket. Finally, we define the Lie transformation associated with the Lie operator $:f_m:$ as

$$\exp(:f_m:) = \sum_{n=0}^{\infty} \frac{:f_m:^n}{n!}. \tag{3.10}$$

It can be shown [2, 3] that the various f_m's occuring in the factored form of $\mathcal{M}$ have a definite physical interpretation for an imaging system. The f_2's represent paraxial

optics, the f_3's are related to second order aberrations, the f_4's to third order aberrations and so on.

We end this section by writing down the symplectic map for two systems which will be used later. First, consider the simple case of transit through a slab of thickness l with refractive index n. It can be easily shown that the symplectic map corresponding to this system is given by

$$\mathcal{M} = \exp[l:(n^2 - p^2)^{1/2}], \tag{3.11}$$

where $p = (p_x, p_y)$. In the factorized form, this is given as

$$\mathcal{M} = \exp\left(-\frac{l}{2n}:p^2:\right)\exp\left(-\frac{l}{8n^3}:(p^2)^2:\right) \times \exp\left(-\frac{l}{16n^5}:(p^2)^3:\right)\exp\left(-\frac{5l}{128n^7}:(p^2)^4:\right)\cdots, \tag{3.12}$$

where $p^2 = p_x^2 + p_y^2$. In order to verify this equation, note that

$$p^f = \mathcal{M}p^i = p^i, \tag{3.13}$$

$$q^f = \mathcal{M}q^i = q^i + l\,p^i/[n^2 - (p^i)^2]^{1/2}, \tag{3.14}$$

where $q = (x, y)$. Using (3.5) in the above equations we get

$$p^f = p^i, \quad q^f = q^i + l(q')^i. \tag{3.15}$$

Remembering that $q' = \mathrm{d}q/\mathrm{d}z$, we see that this is just what would be expected for transit through a distance l along the z-axis.

Note that only even powers of p occur in the expression for $\mathcal{M}$ in (3.12). This is in fact true for all optical systems having the assumed axial and reflection symmetries. These symmetries imply that only the phase space variables p^2, $p \cdot q$ and q^2 (and their various combinations and powers) can occur in f_m's. This automatically implies that only even degree homogeneous polynomials i.e. f_2, f_4, etc. are allowed.

Next, we consider propagation of a light ray through the system given in figure 3. We again assume axial and reflection symmetries. The surface of the interface is parametrized by the equation

$$z^s = \beta_2(x^2+y^2)+\beta_4(x^2+y^2)^2 + \beta_6(x^2+y^2)^3 + \beta_8(x^2+y^2)^4 + \cdots. \tag{3.16}$$

After some computation, one finds the $\mathcal{M}$ (truncated to order 8) for such a surface to be

$$\mathcal{M} = e^{:f_2:}e^{:f_4:}e^{:f_6:}e^{:f_8:}, \tag{3.17}$$

where

$$f_2 = \beta_2(n_1 - n_2)q^2 \tag{3.18}$$

and

$$f_4 = -\beta_2^3\frac{(n_1-n_2)}{n_1}\left[n_1\left(2-\frac{\beta_4}{\beta_2^3}\right) - 2n_2\right](q^2)^2 + 2\beta_2^2\frac{(n_1-n_2)}{n_1}q^2(p\cdot q) + \frac{\beta_2(n_1-n_2)}{2n_1n_2}q^2p^2. \tag{3.19}$$

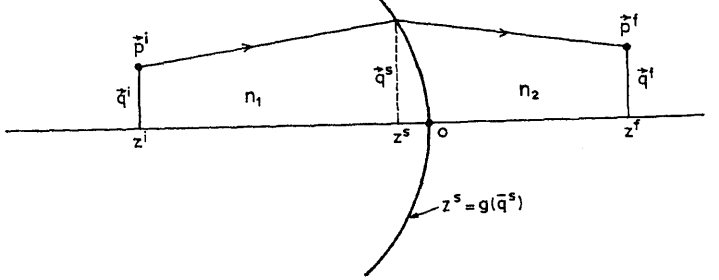


Figure 3. An (axially symmetric) optical system consisting of an interface separating media of refractive indices n_1 and n_2. A ray originating at z^i with initial conditions q^i and p^i passes through the interface and terminates at z^f with final conditions q^f and p^f. The interface is parametrized by the equation $z^s = g(q^s)$.

The expressions for f_6 and f_8 are too long to be documented here. Since we have truncated $\mathcal{M}$ to order 8, the error made in using this truncated expression is also of order 8.

4. Correction of spherical aberration using Lie algebraic method

Before applying the Lie algebraic method to our problem, we briefly describe our plan of action. We will first construct the symplectic map $\mathcal{M}$ (up to order 8) corresponding to our optical system (see figure 2). Since we include all terms up to f_8 in $\mathcal{M}$, we would have taken into account all aberrations up to the seventh order. $\mathcal{M}$ will depend on the surface of the lens and hence on β_2 through β_8 [cf. eq. (3.16)]. Next, we will require that our map satisfy a set of conditions which ensure that $\mathcal{M}$ (and hence our system) is free of spherical aberration up to seventh order. Since $\mathcal{M}$ is a function of β_i's, this would be equivalent to the statement that these β_i's should satisfy certain equations. This will fix our β_i's and hence the equation for the surface of the lens to eight order. Finally, we compare this equation with the exact result obtained in § 2 and show that they agree (up to order 8).

Coming back to our problem, we see that our system (figure 2) has three components:

(1) A transit through distance t in a medium of refractive index n (the lens).
(2) The surface of the lens separating media of refractive indices n and 1 (air) respectively.
(3) A transit through distance l (the focal length) in air ($n = 1$).

The symplectic maps $\mathcal{M}_1$ and $\mathcal{M}_3$ for (1) and (3) are given by expressions of the general form of equation (3.12). The map $\mathcal{M}_2$ for (2) is given by equation (3.17). One of the advantages of the Lie algebraic approach now becomes apparent – the combined map $\mathcal{M}$ for the entire system is just the product of the maps for the individual components:

$$\mathcal{M} = \mathcal{M}_1\mathcal{M}_2\mathcal{M}_3 = \exp\left(-\frac{t}{2n}:p^2:\right)\exp\left(-\frac{t}{8n^3}:(p^2)^2:\right):\cdots$$

$$\exp(\beta_2(n-1):q^2:)\exp(:f_4:)\exp(:f_6:)\exp(:f_8:)\cdots$$

$$\exp\left(-\frac{l}{2}:p^2:\right)\exp\left(-\frac{l}{8}:(p^2)^2:\right)\exp\left(-\frac{l}{16}:(p^2)^3:\right)\exp\left(-\frac{5l}{128}:(p^2)^4:\right)\cdots. \tag{4.1}$$

Using Lie algebraic manipulation techniques [5], we recast the symplectic map $\mathcal{M}$ into the standard form:

$$\mathcal{M} = \exp(:g_2:)\exp(:g_4:)\exp(:g_6:)\exp(:g_8:), \tag{4.2}$$

where g_2, g_4 etc are of the form

$$g_2 = A_2p^2 + B_2p\cdot q + C_2q^2, \tag{4.3}$$

$$g_4 = A_4(p^2)^2 + B_4(p\cdot q)p^2 + \cdots, \tag{4.4}$$

$$g_6 = A_6(p^2)^3 + B_6(p\cdot q)(p^2)^2 + \cdots, \tag{4.5}$$

$$g_8 = A_8(p^2)^4 + B_8(p\cdot q)(p^2)^3 + \cdots. \tag{4.6}$$

Again, the expressions for A_2, B_2 etc have been omitted for the sake of brevity. These coefficients depend on l and n. In addition, the coefficients in g_2 (i.e. A_2, B_2 and C_2) depend on β_2, the coefficients in g_4 on β_2 and β_4 and so on.

Now we are ready to determine β_2, β_4, β_6 and β_8 appearing in the equation for the surface of the lens [cf. eq. (3.16)]. We know that g_2 in $\mathcal{M}$ represents paraxial optics. In the paraxial approximation, it is clear from figure 2 that we want rays that start out parallel (i.e. $w^i = (q,0)$) to come to a focus at a distance l along the z-axis (i.e. $w^f = (0,p)$). Thus,

$$\begin{pmatrix}0\\p\end{pmatrix} = e^{:g_2:}\begin{pmatrix}q\\0\end{pmatrix}. \tag{4.7}$$

Since l and n in g_2 are already given, this equation fixes the only other free parameter β_2 to be

$$\beta_2 = \frac{1}{2l(1-n)}. \tag{4.8}$$

We fix β_4, β_6, and β_8 by imposing conditions which ensure that $\mathcal{M}$ is free of spherical aberration (up to order 7). As defined earlier, spherical aberration is the phenomenon where a given ray is displaced (in the transverse direction) from the paraxial focus on the z-axis by an amount dependent only on its arrival angle (i.e. p in our case).

It can be shown that only the $(p^2)^2$, $(p^2)^3$, and $(p^2)^4$ terms (in g_4, g_6, and g_8 respectively) give rise to spherical aberration of orders 3, 5, and 7 respectively. For example, consider the effect of $(p^2)^2$ term in g_4:

$$q^f = \exp(:A_4(p^2)^2:)q = q + A_4[(p^2)^2, q] = q - 4A_4(p^2)p. \tag{4.9}$$

Thus,

$$\delta q = q^f - q = -4A_4(p^2)p. \tag{4.10}$$

Hence the displacement of the actual arrival point q^f from the Gaussian arrival point q depends only on its Gaussian arrival direction p. This is indeed spherical aberration.

Evidently to eliminate spherical aberration we need to make δq zero. In our case, we can eliminate it up to order 7 by setting

$$A_4 = A_6 = A_8 = 0. \tag{4.11}$$

We can now solve these equations for β_4, β_6, and β_8 in terms of n and l. The resulting expressions (after some manipulation) are

$$\beta_4 = \frac{(n+1)}{8l^3(n-1)^2}, \tag{4.12}$$

$$\beta_6 = -\frac{(n+1)^2}{16l^5(n-1)^3}, \tag{4.13}$$

$$\beta_8 = \frac{5(n+1)^3}{128l^7(n-1)^4}. \tag{4.14}$$

Thus we have succeeded in determining the shape of the lens surface such that our imaging system is free of on-axis spherical aberration (to seventh order). In the $y-z$ plane, this is given by [after substituting the above expressions into eq. (3.16)]

$$z^s = -\frac{y^2}{2l(n-1)} + \frac{(n+1)y^4}{8l^3(n-1)^2} - \frac{(n+1)^2y^6}{16l^5(n-1)^3} + \frac{5(n+1)^3y^8}{128l^7(n-1)^4}. \tag{4.15}$$

We see that this agrees exactly with the equation obtained in § 2 [cf. eq. (2.3)] up to the calculated order.

5. Conclusion

In this paper, we discussed the problem of eliminating on-axis spherical aberration from a point-to-point imaging system. We solved the problem using three methods—direct application of Fermat's principle, ray tracing method, and Lie algebraic method. The results were shown to agree. A direct application of Fermat's principle, even though it gives the exact result in this particular case, is not possible for most optical systems. In contrast, Lie algebraic methods can be applied to a wide variety of optical systems. In particular, they can be applied to systems with no particular symmetry and systems with alignment and positioning errors. They can also be used to analyse graded index systems. All this is possible because the mapping from initial to final conditions in any optical system can be shown [3] to be a symplectic mapping. And symplectic maps can be analysed very efficiently using Lie algebraic methods. In fact, Lie algebraic methods have applications in other fields also [7–10].

Acknowledgements

The authors would like to thank Prof. Alex Dragt for helpful discussions. This work was supported in part by the Department of Science and Technology, Research Grant No. SP/S2/K-28/96.

rences

R K Luneberg, *Mathematical theory of optics* (University of California Press, Berkeley, 1964)
A J Dragt, *J. Opt. Soc. Am.* **72**, 372 (1982)
A J Dragt, E Forest and K B Wolf, in *Lie methods in optics* edited by J S Mondragon and K B Wolf (Springer Verlag, Berlin, 1986)
H Goldstein, *Classical mechanics*, second edition (Addison-Wesley, Reading, 1980)
A J Dragt, F Neri, G Rangarajan, D R Douglas, L M Healy and R D Ryne, *Ann. Rev. Nucl. Part. Sci.* **38**, 455 (1988) and references therein
A J Dragt and E Forest, *J. Math. Phys.* **24**, 2734 (1983)
G Rangarajan, *J. Math. Phys.* **37**, 4514 (1996); *Pramana – J. Phys.* **48**, 129 (1997)
G Rangarajan, *J. Math. Phys.* **38**, 2710 (1997)
F Neri and G Rangarajan, *Phys. Rev. Lett.* **64**, 1073 (1990)
A J Dragt, F Neri and G Rangarajan, *Phys. Rev.* **A45**, 2572 (1992)

PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 645–651

Two body non-leptonic Λ_b decays in the large N_c limit

A K GIRI, L MAHARANA and R MOHANTA
Department of Physics, Utkal University, Bhubaneswar 751 004, India

MS received 10 March 1997; revised 15 September 1997

Abstract. The two body non-leptonic Λ_b decays are analysed in the HQET with factorization approximation and large N_c limit. In this limit, Λ_b and Λ_c baryons can be treated as the bound states of chiral soliton and heavy meson, and consequently the Isgur–Wise function comes out in a straight forward manner. The results obtained remain well below their previously predicted upper limit.

Keywords. Heavy quark effective theory; factorization approximation.

PACS Nos 11.30; 13.30

1. Introduction

In the last few years considerable progress has been achieved in the understanding of the weak decays of heavy hadrons due to the development of the heavy quark effective theory (HQET) [1–5]. Contrary to the significant progress made in the studies of the meson decays, advancement in the arena of heavy baryons has been very slow. In particular the non-leptonic weak decays of heavy baryons have not been understood clearly till now. At present, there are not many experimental results available for heavy baryons. But in the future we may expect more and more data coming from the colliders. Hence the study of heavy baryons is of great interest in the near future. In fact some phenomenological approaches such as pole model, current algebra etc. have been employed to analyse these decay processes. The well known factorization hypothesis [6–12] which has been applied successfully to the heavy meson decays can also be applied for heavy baryon cases for large N_c limit [13].

In this article we intend to study the two body non-leptonic Λ_b decays in the heavy quark effective theory [1–5] considering factorization approximation. The HQET provides a convenient and simplified framework to analyse the weak decays of heavy hadrons, composed of one heavy quark and any number of light quarks. Of particular importance are the semileptonic decays of heavy mesons, where in the limit of infinite quark masses, all the hadronic form factors can be expressed in terms of a single universal function $\xi(v\cdot v')$, the Isgur–Wise function [1]. The function depends only on the four velocities of the heavy particles involved, and is normalized at the point of zero recoil. Similarly in case of weak decays of heavy baryons one can also write the form factors in terms of another Isgur–Wise function [3, 5].

We therefore consider it worthwhile to investigate the two-body nonleptonic decays $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$, where P and V denote pseudoscalar and vector mesons

baryons, in the large N_c limit [13]. The use of HQET implies that the expression for the decay widths should contain the universal Isgur–Wise function, which is normalized at the point of zero recoil. However other than at the point of zero recoil, HQET does not predict the shape of the Isgur–Wise (IW) function. Therefore to know the value of Isgur–Wise function at the particular kinematic point of interest, we evaluate it in the large N_c limit [14, 15], where factorization approximation is valid. Earlier these decays have been studied by Mannel *et al* [13]. They have parametrized the Isgur–Wise form factor $G_1(v\cdot v')$ in three different forms as

$$G_1(v\cdot v') = 1 + \frac{1}{4}a(v-v')^2(v+v')^2, \quad G_1(v\cdot v') = \frac{1}{1-(v-v')^2/\omega_0^2},$$
$$\text{and} \quad G_1(v\cdot v') = \exp[b(v-v')^2], \tag{1}$$

and they have taken the values of the parameters a, b, and ω_0, from the work on B meson decays [16]. They have also predicted the upper limits for the branching ratios for these decay processes, considering the normalized value of the IW function.

The report is organized as follows. In § 2 we present the general framework for the study of the nonleptonic decays in the factorization method. The Isgur–Wise function is evaluated in the large N_c limit in § 3. Section 4 contains results and discussions.

2. General framework

Neglecting the penguin contribution, the four fermion effective Hamiltonian relevant to the $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$ decays is given as [13]

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{UD}^* V_{cb}[C_1(m_b)\mathbf{O}_1 + C_2(m_b)\mathbf{O}_2], \tag{2}$$

with

$$\mathbf{O}_1 = (\bar{D}U)^\mu(\bar{c}b)_\mu \quad \text{and} \quad \mathbf{O}_2 = (\bar{c}U)^\mu(\bar{D}b)_\mu, \tag{3}$$

where G_F is the Fermi coupling constant and the quark current $(\bar{q}'q)_\mu$ is a short hand for $\bar{q}'_\alpha\gamma_\mu(1-\gamma_5)q_\alpha$; α is the color index. U and D are either c, s or u, d quarks. Thus for U, $D = u$, d we have π^- and ρ^- in the final state as P and V while for U, $D = c$, s the final P/V states are D_s/D_s^* mesons. The values of the Wilson coefficients $C_{1,2}$ can be calculated using the leading logarithmic approximation (LLA) [17] and are given as

$$C_1(m_b) = 1.11 \quad \text{and} \quad C_2(m_b) = -0.26. \tag{4}$$

The decays $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$ can occur by the operator $\mathbf{O}_1$ where it is assumed that the $\Lambda_b \to \Lambda_c$ transition is caused by the current operator $(\bar{c}b)$ and that $P^-(V^-)$ are created by the current operator $(\bar{D}U)$. In the factorization approximation it is assumed that the $\Lambda_b \to \Lambda_c$ transition and the $P^-(V^-)$ creation are independent of each other, and hence the amplitude can be written as

$$\langle\Lambda_c^+ P^-(V^-)|(\bar{D}U)(\bar{c}b)|\Lambda_b\rangle = \langle P^-(V^-)|(\bar{D}U)|0\rangle\langle\Lambda_c^+|(\bar{c}b)|\Lambda_b\rangle. \tag{5}$$

e large N_c limit, where factorization approximation is valid, the contribution of $\mathbf{O}_2$ to
decays is suppressed. Therefore one can write the amplitude for the decays
$\to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$ as

$$\mathcal{M}(\Lambda_b \to \Lambda_c^+ P^-(V^-)) = \frac{G_F}{\sqrt{2}} V_{UD}^* V_{cb} C_1(m_b) \langle P^-(V^-)|(\bar{D}U)^\mu|0\rangle \langle \Lambda_c^+|(\bar{c}b)_\mu|\Lambda_b\rangle. \tag{6}$$

valuate the factorized amplitudes we use the following matrix elements:

$$\langle P(p)|(\bar{D}U)^\mu|0\rangle = -if_P p^\mu, \tag{7}$$

$$\langle V(p,\epsilon)|(\bar{D}U)^\mu|0\rangle = f_V M_V \epsilon^\mu, \tag{8}$$

re f_P and f_V are the pseudoscalar and vector meson decay constants respectively. The ix element $\langle \Lambda_c|(\bar{c}b)_\mu|\Lambda_b\rangle$ is given in the HQET [3, 4] as

$$\langle \Lambda_c^+(v',s')|\bar{c}\gamma_\mu(1-\gamma_5)b|\Lambda_b(v,s)\rangle = \eta(v\cdot v')\bar{u}_c(v',s')\gamma_\mu(1-\gamma_5)u_b(v,s), \tag{9}$$

re $\eta(v\cdot v')$ is the baryonic Isgur–Wise function, $u_c(v',s')$ and $u_b(v,s)$ are the spinors e Λ_c and Λ_b baryons. Thus with eqs (6–9) we obtain the decay widths for the decay esses $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$, given as

$$\Gamma(\Lambda_b(v) \to \Lambda_c^+(v')P^-(p)) = \frac{G_F^2}{8\pi M_{\Lambda_b}^2} |V_{UD}^* V_{cb}|^2 C_1^2(m_b) f_P^2 \eta^2(v\cdot v')|\mathbf{p}| \times [(M_{\Lambda_b}^2 - M_{\Lambda_c}^2)^2 - M_P^2(M_{\Lambda_b}^2 + M_{\Lambda_c}^2)], \tag{10}$$

$$\Gamma(\Lambda_b(v) \to \Lambda_c^+(v')V^-(p)) = \frac{G_F^2}{8\pi M_{\Lambda_b}^2} |V_{UD}^* V_{cb}|^2 C_1^2(m_b) f_V^2 \eta^2(v\cdot v')|\mathbf{p}| \times [(M_{\Lambda_b}^2 - M_{\Lambda_c}^2)^2 + M_V^2(M_{\Lambda_b}^2 + M_{\Lambda_c}^2 - 2M_V^2)], \tag{11}$$

re $|\mathbf{p}|$ is the c.o.m momentum of the emitted particles in the rest frame of initial Λ_b on and M's are the corresponding pseudoscalar, vector meson and baryon masses. above expressions for the decay widths contain besides the known quantities, the nown Isgur–Wise function, which can be calculated in the large N_c limit in a simple ner.

valuation of the Isgur–Wise function

e we have presented the evaluation of the Isgur-Wise function in the same manner as gested in ref. [14]. In the large N_c limit the light baryons n, p, Δ etc. can be viewed as ons in the chiral Lagrangian for pion self interaction [18]. The baryons containing a le heavy charm (or bottom) quark are bound states of these solitons with D and D^* (or d B^*) mesons [19–22]. In this paper we use the bound state soliton picture to estimate value of the baryonic Isgur–Wise function. In the ground state of Λ_Q baryons, the light

quarks are in the spin 0 state [25]. Hence in the bound state soliton picture, Λ_Q-type bound state arise when the spin of the light degrees of freedom of the heavy meson and the spin of the nucleon are combined into a spin zero configuration whereas the isospin of the heavy meson and that of the nucleon are combined into an isospin zero state. Other baryons (e.g, the Δ) only contribute to the bound states with higher isospin.

Let the light degrees of freedom of the heavy baryon be denoted by $|I, I_3; s_l, m_l\rangle$, where I and s_l denote their isospin and spin quantum numbers while I_3 and m_l are their third components respectively. Hence the light degrees of freedom of Λ_Q baryon is denoted by $|0,0;0,0\rangle$. The chiral soliton is denoted by $|R,b;R,n)$, where $R = 1/2$ for the nucleon. On the other hand the light degrees of freedom of the heavy meson is given as $|1/2,c;1/2,p\}$.

In the large N_c limit, the binding potential between the chiral soliton and heavy meson is independent of both the isospin and spin of the particles. Hence for the light degrees of freedom of Λ_Q baryon, we have the decomposition as

$$|0,0;0,0(v)\rangle = \int \mathrm{d}^3\mathbf{q}\Phi_Q(\mathbf{q})(1/2,b;1/2,c|0,0)(1/2,n;1/2,p|0,0) \times |1/2,b;1/2,n(-\mathbf{q}+M_B\mathbf{v}))|1/2,c;1/2,p\,(\mathbf{q}+M_H\mathbf{v})\}, \tag{12}$$

where $(j_1,m_1;j_2,m_2|J,M)$'s are the Clebsch–Gordan coefficients. $\Phi_Q(\mathbf{q})$ is the ground state wave function, M_B and M_H are the masses of the chiral soliton and heavy meson respectively.

The spin-1/2 Λ_Q baryon is composed of a spin-1/2 heavy quark and spin-0 light degrees of freedom. Hence the matrix element of the current $\bar{c}\gamma_\mu(1-\gamma_5)b$ between Λ_b and Λ_c baryons is given as

$$\langle\Lambda_c^+(v',s')|\bar{c}\gamma_\mu(1-\gamma_5)b|\Lambda_b(v,s)\rangle = \langle 0,0;0,0(v')|0,0;0,0(v)\rangle\bar{u}_c\gamma_\mu(1-\gamma_5)u_b. \tag{13}$$

Comparing (9) and (13) we obtain the expression for the baryonic IW function as

$$\begin{aligned}\eta(v\cdot v') &= \langle 0,0;0,0(v')|0,0;0,0(v)\rangle \\ &= \int \mathrm{d}^3\mathbf{q}'\mathrm{d}^3\mathbf{q}\Phi_c^*(\mathbf{q}')\Phi_b(\mathbf{q})(1/2,b';1/2,c'|0,0)^* \\ &\quad\times(1/2,n';1/2,p'|0,0)^*(1/2,b;1/2,c|0,0)(1/2,n;1/2,p|0,0) \\ &\quad\times(1/2,b';1/2,n'(-\mathbf{q}'+M_B\mathbf{v}')|1/2,b;1/2,n(-\mathbf{q}+M_B\mathbf{v})) \\ &\quad\times\{1/2,c';1/2,p'(\mathbf{q}'+M_H\mathbf{v}')|1/2,c;1/2,p(\mathbf{q}+M_H\mathbf{v})\}. \end{aligned} \tag{14}$$

Using the normalization conditions for the chiral soliton and heavy meson states, it is found that all the Clebsch–Gordan coefficients in (14) turn out to be unity and the Isgur–Wise function is given as

$$\eta(v\cdot v') = \int \mathrm{d}^3\mathbf{q}\Phi_c^*(\mathbf{q})\Phi_b(\mathbf{q}+M_B(\mathbf{v}-\mathbf{v}')). \tag{15}$$

It is noted from (15) that the IW function depends on the spatial wave function $\Phi_Q(\mathbf{q})$ of the Λ_Q baryon. In the large N_c limit, the binding potential between the heavy meson and

Table 1. Prediction for the branching ratios BR (in %) for the two body nonleptonic Λ_b decays in the large N_c limit.

Decay process	$\|\mathbf{p}\|$ in MeV	$\eta(v\cdot v')$	BR Present calculation	BR Ref. [13]	$BR_{\max}$ Ref. [13]
$\Lambda_b \to \Lambda_c^+ \pi^-$	2355.343	0.456	0.342	$0.46^{+0.20}_{-0.31}$	2.0
$\Lambda_b \to \Lambda_c^+ D_s^-$	1849.748	0.596	1.156	$2.3^{+0.30}_{-0.40}$	6.5
$\Lambda_b \to \Lambda_c^+ \rho^-$	2284.288	0.474	0.954	$0.66^{+0.24}_{-0.40}$	2.5
$\Lambda_b \to \Lambda_c^+ D_s^{*-}$	1765.854	0.621	1.769	$1.73^{+0.20}_{-0.30}$	4.7

he chiral soliton is simple harmonic [15], and hence the wave function is taken as

$$\Phi_Q(\mathbf{q}) = \frac{1}{(\pi^2 \mu_Q \kappa)^{3/8}} \exp\left(-\frac{\mathbf{q}^2}{2\sqrt{\mu_Q \kappa}}\right), \tag{16}$$

where $\mu_Q = M_B M_H/(M_B + M_H)$, is the reduced mass of the bound state, M_H denotes the masses of D/B mesons for Φ_c/Φ_b wave functions. κ is the spring constant and its value is taken to be $(440\,\mathrm{MeV})^3$ [24]. In the rest frame of the initial state, $v = (1, \vec{0})$ and $\mathbf{v}'$ directed along z-axis we obtain the Isgur–Wise function (15) using (16) for non-relativistic recoils i.e., $|\mathbf{v}'|^2 \approx 2(v\cdot v' - 1)$, as

$$\eta(v\cdot v') = \left[\frac{4\sqrt{\mu_b \mu_c}}{(\sqrt{\mu_b} + \sqrt{\mu_c})^2}\right]^{3/4} \exp\left(-\frac{(v\cdot v' - 1)M_B^2}{\sqrt{\kappa}(\sqrt{\mu_b} + \sqrt{\mu_c})}\right). \tag{17}$$

It should be noted from (17) that the Isgur–Wise function slightly deviates from unity at the point of zero recoil. This violation of normalization condition can be explained as follows. The heavy quark symmetry arises in the limit of QCD, where the heavy quark mass m_Q is taken formally to be infinite and in this limit all the hadronic form factors can be expressed in terms of the IW function. However here we have used finite masses for the heavy mesons (i.e., B and D mesons), and hence heavy quark symmetry breaks down. Thus breaking of the heavy flavor symmetry causes a violation of the normalization condition $\eta(1) = 1$.

The product $(v\cdot v')$ is determined by considering the kinematics of the system. Since we are dealing with the two body decays $\Lambda_b(v) \to \Lambda_c^+(v')P^-(p)/V^-(p)$, from momentum conservation we obtain

$$v\cdot v' = \frac{M_{\Lambda_b}^2 + M_{\Lambda_c}^2 - M_{P/V}^2}{2M_{\Lambda_b} M_{\Lambda_c}}. \tag{18}$$

Taking the masses of the particles from [25] the values of the Isgur–Wise functions are calculated with (17) and (18) as presented in table 1.

4. Results and discussion

Having obtained the values of the Isgur–Wise function, we use the following data to estimate the decay widths for the processes $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$ using (10) and

(11). The CKM matrix elements $V_{cb} = 0.041$, $V_{ud} = 0.0976$ and $V_{cs} = 0.9743$ are taken from ref. [25] and the values of the decay constants used are $f_\pi = 130.7\,\text{MeV}$, $f_{D_s} = f_{D_s^*} = 232\,\text{MeV}$ [25] and $f_\rho = 210\,\text{MeV}$ [13]. With these values we have evaluated the branching ratios for several two body non-leptonic Λ_b decays as presented in table 1.

In this work we have estimated the branching ratios for $\Lambda_b \to \Lambda_c^+ P^-$ and $\Lambda_b \to \Lambda_c^+ V^-$ decays in the heavy quark effective theory with factorization approximation. In fact factorization method works well for the description of non-leptonic decays of heavy baryons in the large N_c limit [13]. The use of HQET allows us to write the weak decay form factors in terms of the Isgur–Wise function. However it does not predict the shape of the IW function except at the point of zero recoil, where it is normalized to unity. Therefore to know the value of Isgur–Wise function at the particular kinematic point of interest, we have evaluated it in the large N_c limit [14, 15] considering the bound state soliton picture. These decays have been previously studied in ref. [13] in which they have parametrized the Isgur–Wise function in three different forms (eq. (1)) and used the values of the unknown parameters from the work on B meson decays [16]. However in our case we have evaluated the Isgur–Wise function in the large N_c limit where factorization approximation is valid and the results came out in a straightforward manner. Therefore our predicted results differ from theirs as noted from table 1. They have also predicted the upper limit of the branching fractions for these decays, by considering the normalized value of the Isgur–Wise function. The results of the present investigation lie well below their corresponding upper limit. As the experimental data are expected in the future from the colliders, these results can be verified, which will definitely enrich our understanding in this sector to a greater extent.

Acknowledgement

One of the authors (RM) would like to thank CSIR, Government of India for financial support.

References

[1] N Isgur and M B Wise, *Phys. Lett.* **B232**, 113 (1989); *Phys. Lett.* **B237**, 527 (1990)
[2] A F Falk, H Georgi, B Grinstein and M B Wise, *Nucl. Phys.* **B343**, 1 (1990)
[3] N Isgur and M B Wise, *Nucl. Phys.* **B348**, 276 (1991)
[4] M Neubert, *Phys. Rep.* **245**, 259 (1994)
[5] H Georgi, *Phys. Lett.* **B240**, 447 (1990); *Nucl. Phys.* **B348**, 293 (1991)
[6] M Bauer, B Stech and M Wirbel, *Z. Phys.* **C34**, 103 (1987)
[7] M J Dugan and B Grinstein, *Phys. Lett.* **B255**, 583 (1991)
[8] D Bortoletto and S Stone, *Phys. Rev. Lett.* **65**, 2951 (1990)
[9] J L Rosner, *Phys. Rev.* **D42**, 3732 (1990)
M Tanimoto, K Goda and K Senba, *Phys. Rev.* **D42**, 3741 (1990)
A Acker, S Pakavasa, S F Tuan and S P Rosen, *Phys. Rev.* **D43**, 3083 (1991)
[10] T Mannel, W Roberts and Z Ryzak, *Phys. Lett.* **B259**, 359 (1991)
[11] V Rieckert, *Phys. Rev.* **D47**, 3053 (1993)
[12] A K Giri, L Maharana and R Mohanta, *Pramana – J. Phys.* **48**, 1005 (1997)
[13] T Mannel and W Roberts, *Z. Phys.* **C59**, 179 (1993)
T Mannel, W Roberts and Z Ryzak, *Phys. Lett.* **B259**, 485 (1991)
[14] C K Chow, *Phys. Rev.* **D51**, 1224 (1995)

Jenkins, A Manohar and M B Wise, *Nucl. Phys.* **B396**, 38 (1993)
Mannel, W Roberts and Z Ryzak, *Phys. Lett.* **B248**, 392 (1990); **B254**, 274 (1991)
K Gaillard and B W Lee, *Phys. Rev. Lett.* **33**, 108 (1974)
Altarelli and L Maiani, *Phys. Lett.* **B32**, 351 (1974)
Witten, *Nucl. Phys.* **B223**, 433 (1983)
S Adkins, C R Nappi and E Witten, *Nucl. Phys.* **228**, 552 (1985)
G Callan and I Klebanov, *Nucl. Phys.* **B262**, 365 (1985); *Phys. Lett.* **B202**, 269 (1988)
Rho, D O Riska and N N Scoccola, *Phys. Lett.* **B251**, 597 (1990)
O Riska and N N Scoccola, *Phys. Lett.* **B265**, 188 (1991)
Guralnik, M Luke and A Manohar, *Nucl. Phys.* **B390**, 474 (1993)
Jenkins and A Manohar, *Phys. Lett.* **B294**, 273 (1992)
Jenkins, A Manohar and M B Wise, *Nucl. Phys.* **B396**, 27 (1993)
M Yan, H Y Cheng, C-Y Cheung, G-L Lin, Y C Lin and H-L Yu, *Phys. Rev.* **D46**, 1148
992)
K Chow and M B Wise, *Phys. Rev.* **D50**, 2135 (1994)
rticle Data Group, Review of Particle Properties, *Phys. Rev.* **D54**, part I (1996)

SUBJECT INDEX

Reviews

Cosmological N-body simulations — *J S Bagla and T Padmanabhan* — 161–192

Some aspects of Kaluza–Klein cosmology — *S K Srivastava* — 323–370

General and Mathematical Physics

Differentiability of invariants and Liapunov equations in Hamiltonian systems — *B R Sitaram* — 193–197

Truncated harmonic oscillator and parasupersymmetric quantum mechanics — *B Bagchi, S N Biswas, Avinash Khare and P K Roy* — 199–204

The algebra and geometry of $SU(3)$ matrices — *K S Mallesh and N Mukunda* — 371–383

Accurate energy-levels of the sextic-double-well oscillator $V(X)=X^6-3X^4$ using modified Hill determinant approach — *Biswanath Rath* — 385–388

Hopping of chaotic dynamics in a doubly resonant parametric oscillator — *Jolly Jose, Bimalendu Deb and S Dutta Gupta* — 389–398

The effect of environment induced pure decoherence on the generalized Jaynes–Cummings model — *Gautam Gangopadhyay and S H Lin* — 399–416

Darboux transformation and elementary exact solutions of the Schrödinger equation — *Vladislav G Bagrov and Boris F Samsonov* — 563–580

Detection of determinism and randomness in time series: A method based on phase space overlap of attractors — *M S Gopinathan* — 581–590

Schrödinger picture formalism of Φ^6 theory — *K P Satheesh and K Babu Joseph* — 591–601

Quantum Mechanics

Causal quantum mechanics — *Virendra Singh* — 5–16

A generalized Schrödinger formalism as a Hilbert space representation of a generalized Liouville equation — *R K Varma* — 17–31

The quantum geometric phase as a transformation invariant — *N Mukunda* — 33–40

Supersymmetry in quantum mechanics — *Avinash Khare* — 41–64

Violation of signal locality and unitarity in a merger of quantum mechanics and general relativity — *Partha Ghose* — 65–69

A new class of cosmological models in Lyra geometry
G P Singh and Kalyani Desikan 205–212

Inhomogeneous cosmological models with heat flux
L K Patel, Ramesh Tikekar and Naresh Dadhich 213–224

Isothermal fluid sphere: Uniqueness and conformal mapping
Naresh Dadhich 417–420

Motion of test particles around domain walls
Subenoy Chakraborty and Lalit Biswas 603–607

Einstein pseudotensor and total energy of the universe
N Banerjee and S Sen 609–615

Statistical Physics

Deposition and evaporation of k-mers: Dynamics of a many-state system
Mustansir Barma 155–160

Particle Physics

Quark confinement in $2+1$ dimensional pure Yang–Mills theory
Spenta R Wadia 71–80

CP symmetry and its violation in fundamental interactions
S D Rindani 81–92

Photon production from baryon rich quark-gluon plasma
Pragya, Ashok Goyal and J D Anand 225–230

Some two body non-leptonic B decays and estimation of f_D in the heavy quark effective theory *A K Giri and R Mohanta* 231–238

Investigation of the boundary value problem corresponding to a generalized Skyrme lagrangian *Rajesh Kumar and P K Chattopadhyay* 421–429

Electromagnetic properties of a chiral-plasma medium
H Torres-Silva, C Villarroel, P H Sakanaka and N Reggiani 431–442

Nuclear Physics

Modified WKB transmission for fusion
S V S Sastry and S K Kataria 239–252

Excitation functions of alpha particle induced reactions in aluminium and copper *S Mukherjee, B Bindu Kumar and N L Singh* 253–261

Bulk viscosity of neutron stars
V K Gupta, A Wadhwa, S Singh and J D Anand 443–453

Semi-empirical formula for Λ-binding energies in ground states of light hypernuclei *M Z Rahman Khan, Nasra Neelofer and M A Suhail* 617–622

Measurement of excitation functions and mean projected recoil ranges of nuclei in ^{12}C-induced reactions on vanadium
M Ismail, R P Sharma and M H Rashid 623–633

Atomic and Molecular Physics

Application of density-functional perturbation theory to calculate non-linear polarizabilities of helium-like systems
Arup Banerjee and Manoj K Harbola 455–460

Polarization and dichroism in atomic Auger spectroscopy following absorption of circularly polarized light *S Sen* 461–478

Two Coulomb waves theory for direct excitation
Sima Chakrabarti and Durga P Dewangan 479–491

Plasma Physics

Chaos and its implications *B Buti* 93–107

Nonlinear wave modulations in plasmas *N N Rao* 109–127

Thermodynamics of nonlinear plasma systems *P K Kaw* 129

Lasers, Optics and Spectroscopy

Fine structure in the time of flight distribution of C_2 in laser produced plasma from graphite *S S Harilal, Riju C Issac, C V Bindhu, V P N Nampoori and C P G Vallabhan* 317–322

Instrumentation and Techniques

Measurement of fast neutron induced fission cross-section of thorium using Lexan plastic track detector
R K Jain, H S Virk, J Rama Rao and S K Bose 515–519

SCORPION: A system for coincidences between recoil and projectile ions at NSC, New Delhi *M J Singh, S K Goel, R Shanker, D O Kataria, N Madhavan, P Sugathan, J J Das, D K Avasthi and A K Sinha* 521–533

TDPAC measurements in textured hcp samples
R Govindaraj and K P Gopinathan 535–546

Condensed Matter Physics

Projection method for 4d non-orthogonal hyperlattices
S Jayanthi, R K Mandal and S Lele 263–267

Four constant potential function of amorphous selenium and evaluaiton of its thermodynamic and acoustic properties
R Venkatesh, R V Gopala Rao and Sumita Bandyopadhyay 269–283

High pressure study of phase transitions in α-$FePO_4$
Chitra Murli, Surinder M Sharma, S K Kulshrestha and S K Sikka 285–291

Thermoelectric power of high T_c superconductors using extended Hubbard model *S Keshri and P K Barhai* 293–300

Radially dependent antishielding factor $\gamma(r)$ of the Sc^{3+} ion in the solid state *B B Panigrahi, N C Mohapatra and S Hafizuddin* 301–308

Time dependent magnetization of a $Zn_{0.1}Fe_{0.9}Fe_2O_4$ fine magnetic particle system *R V Upadhyay* 309–316

On the water-induced critical double point in a polymer solution: The role of isotope substitution *B M Jaffar Ali, S Venkatachalam and A Kumar* 493–504

Self-consistent numerical study of pure and impurity doped three-band Hubbard model *Prasenjit Sen and Avinash Singh* 505–514

Astrophysics

The large scale structure of the universe: Dynamical and statistical aspects *Varun Sahni* 131–145

Rotation curves of galaxies: Missing mass or missing physics *V Krishan* 147–154

Brief Reports

Phonon dispersion in aluminium arsenide and antimonide *Prafulla K Jha, Sujata Rath and Sankar P Sanyal* 547–553

Slowness surfaces, ray velocity surfaces and phonon focussing in high T_c superconductor crystals *J Philip and M S Kala* 555–561

Spherical aberration and its correction using Lie algebraic methods *Govindan Rangarajan and Minita Sachidanand* 635–643

Two body non-leptonic Λ_b decays in the large N_c limit *A K Giri, L Maharana and R Mohanta* 645–651

AUTHOR INDEX

Anand J D
see Pragya 225
see Gupta V K 443
Avasthi D K
see Singh M J 521

Babu Joseph K
see Satheesh K P 591
Bagchi B
Truncated harmonic oscillator and parasupersymmetric quantum mechanics 199
Bagla J S
Cosmological N-body simulations 161
Bagrov Vladislav G
Darboux transformation and elementary exact solutions of the Schrödinger equaton 563
Bandyopadhyay Sumita
see Venkatesh R 269
Banerjee Arup
Application of density-functional perturbation theory to calculate nonlinear polarizabilities of helium-like systems 455
Banerjee N
Einstein pseudotensor and total energy of the universe 609
Barhai P K
see Keshri S 293
Barma Mustansir
Deposition and evaporation of k-mers: Dynamics of a many-state system 155
Bindhu C V
see Harilal S S 317
Biswas Lalit
see Chakraborty Subenoy 603
Biswas S N
see Bagchi B 199
Bose S K
see Jain R K 515
Buti B
Chaos and its implications 93

Chakrabarti Sima
Two Coulomb waves theory for direct excitation 479
Chakraborty Subenoy
Motion of test particles around domain walls 603
Chattopadhyay P K
see Kumar Rajesh 421
Dadhich Naresh
see Patel L K 213
Isothermal fluid sphere: Uniqueness and conformal mapping 417
Das J J
see Singh M J 521
Deb Bimalendu
see Jose Jolly 389
Desikan Kalyani
see Singh G P 205
Dewangan Durga P
see Chakrabarti Sima 479

Gangopadhyay Gautam
The effect of environment induced pure decoherence on the generalized Jaynes–Cummings model 399
Ghose Partha
Violation of signal locality and unitarity in a merger of quantum mechanics and general relativity 65
Giri A K
Some two body non-leptonic B decays and estimation of f_D in the heavy quark effective theory 231
Two body non-leptonic Λ_b decays in the large N_c limit 645
Goel S K
see Singh M J 521
Gopinathan K P
see Govindaraj R 535
Gopinathan M S
Detection of determinism and randomness in time series: A method based on phase space overlap of attractors 581
Govindaraj R
TDPAC measurements in textured hcp samples 535

Goyal Ashok
see Pragya 225
Gupta S Dutta
see Jose Jolly 389
Gupta V K
Bulk viscosity of neutron stars 443

Hafizuddin S
see Panigrahi B B 301
Harbola Manoj K
see Banerjee Arup 455
Harilal S S
Fine structure in the time of flight distribution of C_2 in laser produced plasma from graphite 317

Ismail M
Measurement of excitation functions and mean projected recoil ranges of nuclei in ^{12}C-induced reactions on vanadium 623
Issac Riju C
see Harilal S S 317

Jaffar Ali B M
On the water-induced critical double point in a polymer solution: The role of isotope substitution 493
Jain R K
Measurement of fast neutron induced fission cross-section of thorium using Lexan plastic track detector 515
Jayanthi S
Projection method for 4d non-orthogonal hyperlattices 263
Jha Prafulla K
Phonon dispersion in aluminium arsenide and antimonide 547
Jose Jolly
Hopping of chaotic dynamics in a doubly resonant parametric oscillator 389

Kala M S
see Philip J 555
Kataria D O
see Singh M J 521
Kataria S K
see Sastry S V S 239
Kaw P K
Thermodynamics of nonlinear plasma systems 129
Keshri S
Thermoelectric power of high T_c superconductors using extended Hubbard model 293
Khare Avinash
see Bagchi B 199
Supersymmetry in quantum mechanics 41
Krishan V
Rotation curves of galaxies: Missing mass or missing physics 147
Kulshreshtha S K
see Murli Chitra 285
Kumar A
see Jaffar Ali B M 493
Kumar Bindu B
see Mukherjee S 253
Kumar Rajesh
Investigation of the boundary value problem corresponding to a generalized Skyrme lagrangian 421
Lele S
see Jayanthi S 263

Lin S H
see Gangopadhyay Gautam 399

Madhavan N
see Singh M J 521
Maharana L
see Giri A K 645
Mallesh K S
The algebra and geometry of $SU(3)$ matrices 371
Mandal R K
see Jayanthi S 263
Mohanta R
see Giri A K 231, 645
Mohapatra N C
see Panigrahi B B 301
Mukherjee S
Excitation functions of alpha particle induced reactions in aluminium and copper 253
Mukunda N
The quantum geometric phase as a transformation invariant 33
see Mallesh K S 371
Murli Chitra
High pressure study of phase transitions in α-$FePO_4$ 285

Nampoori V P N
see Harilal S S 317
Neelofer Nasra
see Rahman Khan M Z 617

Padmanabhan T
see Bagla J S 161
Panigrahi B B
Radially dependent antishielding factor $\gamma(r)$ of the Sc^{3+} ion in the solid state 301
Patel L K
Inhomogeneous cosmological models with heat flux 213
Philip J
Slowness surfaces, ray velocity surfaces and phonon focussing in high T_c superconductor crystals 555
Pragya
Photon production from baryon rich quark-gluon plasma 225

Rahman Khan M Z
Semi-empirical formula for Λ-binding energies in ground states of light hypernuclei 617
Rangarajan Govindan
Spherical aberration and its correction using Lie algebraic methods 635
Rao J Rama
see Jain R K 515

Rashid M H
see Ismail M 623
Rath Biswanath
Accurate energy-levels of the sextic-double-well oscillator $V(X) = X^6 - 3X^4$ using modified Hill determinant approach 385
Rath Sujata
see Jha Prafulla K 547
Reggiani N
see Torres-Silva H 431
Rindani S D
CP symmetry and its violation in fundamental interactions 81
Roy P K
see Bagchi B 199

Sachidanand Minita
see Rangarajan Govindan 635
Sahni Varun
The large scale structure of the universe: Dynamical and statistical aspects 131
Sakanaka P H
see Torres-Silva H 431
Samsonov Boris F
see Bagrov Vladislav G 563
Sanyal Sankar P
see Jha Prafulla K 547
Sastry S V S
Modified WKB transmission for fusion 239
Satheesh K P
Schrödinger picture formalism of Φ^6 theory 591
Sen Prasenjit
Self-consistent numerical study of pure and impurity doped three-band Hubbard model 505
Sen S
Polarization and dichroism in atomic Auger spectroscopy following absorption of circularly polarized light 461
see Banerjee N 609
Shanker R
see Singh M J 521
Sharma R P
see Ismail M 623
Sharma Surinder M
see Murli Chitra 285
Sikka S K
see Murli Chitra 285
Singh Avinash
see Sen Prasenjit 505
Singh G P
A new class of cosmological models in Lyra geometry 205
Singh N L
see Mukherjee S 253
Singh S
see Gupta V K 443
Singh Virendra
Causal quantum mechanics 5
Sinha A K
see Singh M J 521
Sitaram B R
Differentiability of invariants and Liapunov equations in Hamiltonian systems 193
Srivastava S K
Some aspects of Kaluza–Klein cosmology 323
Sugathan P
see Singh M J 521
Suhail M A
see Rahman Khan M Z 617

Tikekar Ramesh
see Patel L K 213
Torres-Silva H
Electromagnetic properties of a chiral-plasma medium 431

Upadhyay R V
Time dependent magnetization of a $Zn_{0.1}Fe_{0.9}Fe_2O_4$ fine magnetic particle system 309

Vallabhan C P G
see Harilal S S 317
Varma R K
A generalized Schrödinger formalism as a Hilbert space representation of a generalized Liouville equation 17
Venkatachalam S
see Jaffar Ali B M 493
Venkatesh R
Four constant potential function of amorphous selenium and evaluation of its thermodynamic and acoustic properties 269
Villarroel C
see Torres-Silva H 431
Virk H S
see Jain R K 515

Wadhwa A
see Gupta V K 443
Wadia Spenta R
Quark confinement in $2+1$ dimensional pure Yang–Mills theory 71

End of the forty ninth volume